AOC10-1

POPQAE 50 (1) 143–284 (1986)
ISSN 0033-362X

Public Opinion Quarterly

Homicide and Bargained Justice: The Agenda-Setting Effect of Crime News on Prosecutors *David Pritchard*

Interpersonal Communication and News Comprehension *John P. Robinson and Mark R. Levy*

Measuring Change in Personal Economic Well-Being *Steven J. Rosenstone, John Mark Hansen, and Donald R. Kinder*

Race Differences in Abortion Attitudes *Elaine J. Hall and Myra Marx Ferree*

Picking the Winners: Politician vs. Voter Predictions of Two Controversial Ballot Measures *James B. Lemert*

Election Predictions: An Empirical Assessment *William Buchanan*

The Presence of Others and Overreporting of Voting in American National Elections *Brian D. Silver, Paul R. Abramson, and Barbara A. Anderson*

Opinions on Fictitious Issues: The Pressure to Answer Survey Questions *George F. Bishop, Alfred J. Tuchfarber, and Robert W. Oldendick*

Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys *Robert M. Groves and Lou J. Magilavy*

An Experiment with a Confidentiality Reminder in a Telephone Survey *James H. Frey*

THE POLLS: Environmental Protection *John M. Gillroy and Robert Y. Shapiro*

BOOK REVIEWS

SUMMER 1986

Published for The American Association for Public Opinion Research
by The University of Chicago Press

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR ADVISORY COMMITTEE

LEO BOGART BURNS W. ROPER
NORMAN M. BRADBURN BERNARD ROSHCO
ANDREW KOHUT AND ELEANOR SINGER, COUNCILLORS-AT-LARGE

Editor: ELEANOR SINGER *Managing Editor:* INA C. JAEGER
Department Editors:
PHILIP A. HARDING, *News and Notes*
MARC B. GLASSMAN, *Book Reviews*
TOM W. SMITH, *The Polls*

EDITORIAL BOARD

JAMES R. BENIGER
JOAN S. BLACK
NORMAN M. BRADBURN
CHARLES F. CANNELL
MURIEL G. CANTOR
ALBERT H. CANTRIL
DONALD DILLMAN
MARTIN R. FRANKEL
HERBERT J. GANS
ALBERT E. GOLLIN
ROBERT HORNIK
MARY R. JACKMAN
F. THOMAS JUSTER
RONALD C. KESSLER
MARK R. LEVY
PHILIP MEYER
J. RONALD MILAVSKY
WARREN J. MITOFSKY
ELISABETH NOELLE-NEUMANN
STANLEY PRESSER
DONALD F. ROBERTS
MORRIS ROSENBERG
CHARLES R. WRIGHT
DANIEL YANKELOVICH

Volume #50, 4 issues (1986)

© 1986 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, Chicago. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $38.00; individuals 1 year $19.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $17.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $17.00. Other countries add $3.00 for each year's subscription to cover postage.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Journalism Building, Columbia University
116th Street and Broadway, New York, New York 10027
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

Public Opinion Quarterly

VOLUME 50 SUMMER 1986 NUMBER 2

Homicide and Bargained Justice: The Agenda-Setting Effect of Crime News on Prosecutors *David Pritchard* 143

Interpersonal Communication and News Comprehension *John P. Robinson and Mark R. Levy* 160

Measuring Change in Personal Economic Well-Being *Steven J. Rosenstone, John Mark Hansen, and Donald R. Kinder* 176

Race Differences in Abortion Attitudes *Elaine J. Hall and Myra Marx Ferree* 193

Picking the Winners: Politician vs. Voter Predictions of Two Controversial Ballot Measures *James B. Lemert* 208

Election Predictions: An Empirical Assessment *William Buchanan* 222

The Presence of Others and Overreporting of Voting in American National Elections *Brian D. Silver, Paul R. Abramson, and Barbara A. Anderson* 228

Opinions on Fictitious Issues: The Pressure to Answer Survey Questions *George F. Bishop, Alfred J. Tuchfarber, and Robert W. Oldendick* 240

Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys *Robert M. Groves and Lou J. Magilavy* 251

An Experiment with a Confidentiality Reminder in a Telephone Survey *James H. Frey* 267

THE POLLS: Environmental Protection *John M. Gillroy and Robert Y. Shapiro* 270

BOOK REVIEWS *Marc B. Glassman, Editor* 280

Sandra J. Ball-Rokeach, Milton Rokeach, and Joel W. Grube, *The Great American Values Test: Influencing Behavior and Belief Through Television* *Melvin M. Mark* 280

Willard D. Rowland, Jr. and Bruce Watkins, EDS., *Interpreting Television: Current Research Perspectives* *Muriel G. Cantor* 282

NOTICE TO CONTRIBUTORS

Three copies of each article should be submitted. Authors are urged not to submit original typescripts. Copies of manuscripts which are reviewed become the property of the journal, and unused copies will be returned only if the author provides a stamped, self-addressed envelope. Bibliographic-references, footnotes, and tables should be typed to conform to the printed style of this journal.

All submissions that make use of survey data not archived at the Inter-University Consortium for Political and Social Research or the Roper Center must contain the following information: (1) A definition of the population from which the sample was drawn and a statement of the selection procedure sufficient to permit at least approximate replication. (2) The response rate and details of its calculation. (If response rate is not appropriate, the number of refusals.) (3) Dates the survey was conducted. (4) Exact wording of all questions used in the analysis.

Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright in the article to the publisher. This transfer will insure the widest possible dissemination of information under the U.S. Copyright Law.

PUBLIC OPINION QUARTERLY is hospitable to all points of view, provided that the material presented will help to illuminate problems of communication and public opinion. The American Association for Public Opinion Research assumes responsibility only for material signed by an officer of the Association in his or her official capacity. Statements of fact and opinion are made on the responsibility of the authors alone and do not represent the views of the editors.

PUBLIC OPINION QUARTERLY has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by the U.S. Copyright Law. If no code appears on an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1978 may be copied for a per-copy fee of $1.00, also payable through the Center. The consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotional purposes, or for creating new collective works.

Homicide and Bargained Justice: The Agenda-Setting Effect of Crime News on Prosecutors

DAVID PRITCHARD

THE EFFECT of publicity upon the adversary system of criminal justice has received an enormous amount of attention in recent years. The Supreme Court of the United States has addressed the fair trial/free press issue several times, and many social scientists have studied the effect of publicity on potential jurors (for reviews of the empirical research see Bush, 1970; Connors, 1975; Simon, 1977; Buddenbaum et al., 1981).

Generally overlooked in the fair trial/free press debate, however, is the fact that as many as 90 percent of all criminal convictions in the United States are the result of plea bargaining rather than full-blown adversary trials (Heumann, 1978; Brosi, 1979). The defendant admits guilt in return for some implicit or explicit concession from the prosecution. Because there is no jury, there is no chance that press coverage will prejudice the jury.

Nonetheless, it is possible that press coverage will taint the *process*. This study addresses that issue by examining the relationship between

---

Abstract This article reports the findings of research into the possibility that differences in newspaper coverage of individual criminal cases may influence the behavior of key justice-system officials with respect to those cases. The study analyzed police and court records regarding all people arrested for homicide over an 18-month period in Milwaukee County, Wisconsin. Newspaper coverage of those cases also was analyzed. The results suggest that the amount of space newspapers devote to a criminal case helps set the agenda of at least one class of public officials—prosecutors who must decide which criminal cases to plea-bargain, and which to take to trial.

David Pritchard is Assistant Professor at the Indiana University School of Journalism, Bloomington, Indiana 47405. An earlier version of this article was presented to the Annual Meeting of the Law and Society Association, San Diego, California, in June 1985. The author wishes to thank Professors William A. Hachten, Robert E. Drechsel, Sharon Dunwoody, and Stewart Macaulay for insightful comments on previous drafts.

Public Opinion Quarterly Vol. 50:143–159 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-143/$2.50

newspaper coverage of individual cases and whether prosecutors engage in plea bargaining in those cases.

It is an important issue to consider, because the results of the consensual process epitomized by plea bargaining are quite different from those of the adversarial process epitomized by jury trials. For example, defendants found guilty by a jury are more likely to be sentenced to a period of incarceration, and can expect to receive substantially longer sentences than those whose cases are plea bargained, everything else being equal (Uhlman and Walker, 1980; Shane-DuBow et al., 1981; Brereton and Casper, 1982; Pruitt and Wilson, 1983).

Previous Work

Prosecutors are avid readers of newspaper stories about their cases, and most say that the news media are good indicators of the public image of the criminal justice system (Drechsel, 1983). So it is not entirely surprising that some prosecutors acknowledge that they take press coverage of a case into consideration in deciding whether to engage in plea negotiations.

For example, a study of plea bargaining in two California jurisdictions found that prosecutors and defense lawyers were less likely to agree in negotiations if a case had received news coverage (Utz, 1978). Surveys of prosecutors and public defenders in Cook County, Illinois, also suggested that a substantial number of prosecutors would not plea bargain in cases that had received news coverage (Jones, 1978). One Cook County prosecutor said: "I'd be crazy to bargain in any case that received a lot of news coverage. The public will demand justice, which means that I usually go to trial" (Jones, 1978:200). A Cook County public defender claimed that prosecutors refused to negotiate in publicized cases because they are "afraid of being considered soft on criminals" (Jones, 1978:202; see also Kaplan, 1965; Miller, 1970; and Eisenstein and Jacob, 1977).

Why might publicity make a prosecutor unwilling to plea bargain a case? Alschuler (1968) studied prosecutors in a dozen large American cities, and found that they were motivated more by what they perceived to be their self-interest than by considerations of justice or fairness for the defendant.

> It is politically embarrassing for a prosecutor to consent to a light sentence for a serious offender, even as part of a plea agreement. It is usually easier, in fact, for a prosecutor to explain an acquittal than it is for him to explain a plea agreement in a publicized case; an acquittal can be attributed to the finder of fact. In routine cases, by contrast, there is little danger that sentencing will

> attract public attention. A prosecutor can maximize the political value of these cases by securing the maximum number of convictions, regardless of the price he pays in sentencing concessions (Alschuler, 1968:107).

In other words, prosecution is a political process, and prosecutors have a political stake in how their actions are perceived. Maintaining a public image as a crime-fighter is important to the prosecutor, perhaps to the extent of stressing adversary dispositions in publicized cases, regardless of the strength of evidence against the defendant. That way, if something "good" (like a conviction) happens, the prosecutor can take the credit. If something "bad" (like an acquittal or dismissal of charges) happens, the prosecutor can implicitly or explicitly shift the responsibility to the judge or to the jury. The result is that the blame is transferred (Newman, 1966; Galanter, 1983). Prosecutors may be especially likely to act this way in homicide cases (Alschuler, 1968).

In some cases, however, prosecutorial self-interest can lead to negotiations. This is most likely to happen when the prosecutor feels a need to get a conviction—any kind of conviction—despite weak evidence against the defendant. Alschuler writes: "Political considerations may, on occasion, make it important for a prosecutor to secure a conviction for a particular crime, and plea negotiations may provide the only practical means of achieving this objective" (Alschuler, 1968:109).

However, prosecutors' most common reaction to publicity, the existing research makes clear, is a desire to avoid being perceived as soft on criminals.

The notion that the press may influence prosecutorial decision making has clear parallels with the familiar agenda-setting hypothesis, which suggests that the relative prominence of an issue in the news media will influence how salient that issue is to members of the audience (McCombs and Shaw, 1972; McLeod et al., 1974; Becker et al., 1975; Shaw and McCombs, 1977).

Most agenda-setting research has focused on possible effects on ordinary citizens. Only a few studies have explicitly tested the hypothesis that the press may help set the agendas of public officials. Gormley (1975) found mixed evidence for an agenda-setting effect of newspapers on state legislators. Lambeth (1978) concluded that the press helped set energy policy makers' agendas. In a study notable for the fact that it used a direct, rather than a self-reported measure of public-official behavior, Gilberg et al. (1980) found that press content set the agenda for Jimmy Carter's 1978 State of the Union speech. Swank et al. (1982) suggested, but did not test, the hypothesis that newspapers contribute to the salience of crime on local political agendas. Cook et al. (1983) found that a televised investigative report changed the agenda not only of citizens but also of elected officials.

On the other hand, Walker (1977) concluded that for three safety-related issues, the agenda of the U.S. Senate set the agenda for the *New York Times*. And Protess et al. (1985) found that the only agenda affected by a newspaper investigative series was future press coverage of the subject of the series.

In addition to the agenda-setting literature, research into reporter-source interactions contains considerable speculation (and some anecdotal evidence) suggesting that journalists' decisions about which stories to play up can influence the behavior of public officials (see, e.g., Matthews, 1960; Cohen, 1963; Dunn, 1969; Sigal, 1973; Weiss, 1974; Miller, 1978; and Peters, 1980).

Despite these occasional forays, how the news media affect the behavior of public officials is an "inadequately researched area" (Rivers et al., 1975:223). Rivers et al. suggested that research should attempt to determine under what conditions publicity alters "personal and departmental resources in ways that affect the decision-making process" (1975:231).

Such research is especially important in the criminal justice area. Stanga noted that "the influence the press has, if any, on the prosecutor is a neglected area of inquiry" (1971:264). More recently, Buddenbaum et al. came to a similar conclusion, suggesting that "a study of the effects of pretrial publicity on the disposition of cases might also be useful, to establish how much influence pretrial publicity has on the way cases are handled and on the amount of time and money spent for similar cases that receive varying publicity" (1981:7).

The Context of the Study

This study focuses on prosecutors in the district attorney's office in Milwaukee County, Wisconsin. The basic business of prosecutorial organizations is prosecuting criminal cases. In Milwaukee, as elsewhere in the United States, such cases are processed primarily by plea bargaining, which serves the cause of conserving relatively scarce organizational resources (Brosi, 1979). At the same time, however, it is not a popular way of settling cases (Hearst Corporation, 1983). In Milwaukee, for example, less than a quarter of the population favors plea bargaining (Metropolitan Milwaukee Criminal Justice Council, 1980). In addition, one of the city's newspapers flatly opposes the practice (Wills, 1977), while the other is merely skeptical (Milwaukee Journal, 1981).

Although in theory defendants decide whether to plead guilty or to exercise their constitutional right to a jury trial, in practice the prosecutor generally controls whether a case is plea bargained (Blumberg, 1967; Alschuler, 1968; Casper, 1972; Alschuler, 1975; Heumann, 1978; Gifford, 1981; Gifford, 1983). On the one hand, limited resources compel prosecutors to use plea bargaining as the normal mode of settling criminal cases. On the other hand, the public's anger about crime suggests that the electorate may not respond warmly to a prosecutor who plea bargains a visible case.

Clearly, then, not all cases can go to full-blown adversarial trial (resource constraints), but some trials are necessary (the influence of public opinion).

This study hypothesizes that newspaper coverage of a criminal case influences whether prosecutors engage in plea bargaining in a given case, and that the more extensive the newspaper coverage of a case, the less likely the district attorney's office is to negotiate in the case.

Data

Data to test this study's hypothesis were extracted from police and court records and from news stories. Information was obtained on every nonvehicular homicide case presented to the district attorney's office for possible prosecution during the 18-month period between January 1, 1981 and June 30, 1982. The study focuses on homicides because lesser crimes seldom receive press coverage in a major metropolitan community like Milwaukee. Three homicides that had not been disposed of by May 31, 1983 were eliminated from the analysis. Prosecutors filed homicide charges against every suspect arrested for homicide during the study period. In all, the cases of 90 homicide defendants were included in this study.

To find out how the cases were processed, every publicly available document on each case was scrutinized. Included were the inmate registration log at the Milwaukee County Jail; case files at the office of the Milwaukee County Clerk of Courts, Felony Division, which contained copies of criminal complaints, autopsy reports and other pieces of documentary evidence, summaries (and sometimes transcripts) of hearings, and memos from the prosecution and defense; and all news items about the cases published by Milwaukee's daily newspapers, both owned by the Journal Company, which granted full access to its files of clippings. The editions of the morning *Sentinel* and the evening and Sunday *Journal* that circulate in Milwaukee County carried 744 staff-written items—news stories, editorials, and columns of staff opinion—about the homicides in this study.

Measures

The independent variable in this study is the level of newspaper interest in the defendant's case. The dependent variable is the behavior of the prosecutor's office with respect to the case.

LEVEL OF NEWSPAPER INTEREST IN A CASE

Conceptually, this variable is a function of the resources the newspaper organization is willing to devote to a case. Although newspapers have at their disposal several kinds of resources, the two that are most visible to outsiders like prosecutors are space in the paper and staff effort.

The amount of space a newspaper devotes to a case can be measured in a straightforward manner. The number of news items about a case is an interval-level variable. So is the number of paragraphs written about a case. This study divides total number of paragraphs by total number of stories to come up with a measure of the level of newspaper interest in the case: the average length of news items about the case, excluding editorials (which tend to be quite short) and excluding stories published after the prosecutor's discretionary period.

That discretionary period is defined as lasting from the date of the homicide until the mode of disposition was known, generally at the beginning of a trial or the acceptance of a guilty plea by a judge. Ending the discretionary period any earlier would skew the results of the analysis, because plea agreements can be—and in several cases were—made or unmade at the last minute. Cases that seem irrevocably headed toward adversary dispositions sometimes are negotiated on the eve of a scheduled trial. Similarly, defendants may accept plea bargains, but then change their minds at the hearing and refuse to plead guilty.

Average story length, rather than total number of stories or paragraphs, is used because average story length in theory is not a function of how long a case lingers in the felony disposition process. The total number of stories or paragraphs the press devotes to a case, on the other hand, can be more a function of how many pretrial hearings take place than of how interested the press is in the case.

Number of paragraphs, rather than number of column inches, is used to measure story length because varying column sizes in both newspapers made computing standardized column inches very difficult. Standardized column inches and number of paragraphs are highly correlated, however (Budd, 1964).

Both Milwaukee newspapers covered homicides similarly (Pritchard, 1985). For that reason, this study combines the newspapers' coverage to form a single variable: the average length of the stories the Milwaukee newspapers published about a defendant's case. Separate analyses were conducted for each newspaper's coverage, with results virtually identical to the result produced by the combined coverage variable.

The newspapers published no stories at all about 5 of the 90 homicide prosecutions in this study. Those cases are coded as having an average story length of 0.

The amount of newspaper staff effort devoted to a given case was measured by analyzing newspaper content. The typical piece of crime news comes directly from routine law enforcement or judicial sources, often documentary sources such as police blotters, jail logs, and criminal complaints (Stanga, 1971; Cohen, 1975; Sherizen, 1978; Drechsel et al., 1980). Reporters who cover police and court beats can gather such news with relatively little effort.

On some cases, however, reporters do more. They may use nonroutine sources such as witnesses to the crime, friends and relatives of the suspect and/or victim, or not-for-attribution comments from law enforcement officials. Editorials and staff-written columns of opinion also represent nonroutine kinds of coverage.

Accordingly, stories about homicide cases can be categorized either as "routine" (if only routine sources were used) or "nonroutine" (if at least one nonroutine source was used). Of the 744 news items in this study, 13.4 percent were nonroutine by this definition.[1] The variable created using the routine/nonroutine dichotomy is the proportion of nonroutine stories in the newspapers' coverage of a case.

This study did not attempt to determine whether predisposition coverage of individual cases was fair or accurate, however those terms might be defined. The independent variables attempt only to operationalize the level of press interest in a case, measured by the amount of organizational resources newspapers devoted to covering the case.

PROSECUTOR'S PLEA BARGAINING BEHAVIOR

Plea bargaining in this study is measured not by whether a case was settled consensually, but by whether the prosecutor actually engaged in plea negotiations.This "negotiated/did-not-negotiate" variable measures the prosecutor's actual behavior, not whether that behavior led to a consensual outcome. In theory, the negotiated/did-not-negotiate variable is untainted by the behavior of other actors in the criminal justice system. For example, although negotiations cannot succeed without the participation of the prosecutor, a willing prosecutor can be stymied by a defense lawyer or defendant unwilling to accept the prosecutor's terms. Such cases would be disposed of adversarially, but not because the prosecutor failed to negotiate.

---

[1] All coding was done by the author. To test intercoder reliability, a colleague coded a 10 percent sample of the 744 news items. There was agreement on 96 percent of the items, resulting in a reliability coefficient (Scott's pi) of .85 (Scott, 1955).

Court records and/or news items contained explicit evidence that prosecutors engaged in negotiations in 45 of the 90 cases in this study. In 35 of the cases in which prosecutors negotiated, the result was a consensual settlement, in which the prosecutor and the defense agreed on the appropriate disposition of the case. In the remaining 10 cases, defendants refused plea bargains offered by prosecutors. Those cases either went to jury trial or were dismissed over the objection of the prosecution.

CONTROL VARIABLES

One of the most challenging tasks in a study of this nature is to hold constant factors that may affect both the independent and the dependent variables. In an attempt to achieve that goal, this study uses an extensive set of control variables. Much of the variation in crime seriousness, a factor that could influence both newspaper interest and prosecutors' plea bargaining behavior, is implicitly controlled by this study's exclusive focus on homicides.

Homicides themselves can vary in a number of ways, however, so additional control variables are used. They include personal attributes (race, age, and sex) of the homicide suspect and victim; whether the suspect and the victim knew each other; the suspect's prior record; the initial charge against the defendant (first-degree murder or a lesser homicide charge, such as manslaughter); whether the defendant was charged with crimes beyond the first homicide count; and the number of suspects alleged to have been involved in the homicide.

Some factors, such as the prominence of the homicide suspect and victim or the bizarreness of the incident, are difficult to quantify. In most cases, however, court records and news stories contain enough details of the incident and of the people involved in it to permit a qualitative evaluation of such aspects of the case.

Results

Discriminant analysis (Cohen and Cohen, 1975; Klecka, 1975) was used to test the hypothesis concerning the influence of newspaper coverage on plea bargaining behavior. The analysis supported the hy-

Table 1. Results of Discriminant Analysis, with Whether the Prosecution Negotiated as the Dependent Variable

Canonical correlation squared	.238
Improvement in ability to predict	53.3%
Relative contribution of significant discriminating variables:	
Average story length	34.8%
Suspect knew victim	26.9%
Prior record	19.5%
Initial charge	12.4%
Multiple charges	6.5%

pothesis: press behavior—specifically, the average length of stories about a case—was the strongest predictor of whether prosecutors engaged in negotiations. The proportion of stories about a case that relied partly or entirely on nonroutine sources, however, was not a significant predictor of negotiations.

Table 1 shows the results of the analysis. Of all the variables, five proved to be statistically significant (at $p = .05$) predictors: average story length (the shorter the average story, the more likely the prosecutor would negotiate); whether the defendant and the victim knew each other (negotiations were more likely if the defendant and victim had been acquainted); the defendant's prior record (negotiations were more likely if the defendant had no prior record); the seriousness of the initial charge (negotiations were more likely when charges were less serious); and whether the defendant faced multiple charges stemming from the incident (negotiations were more likely when there was only one charge).

The canonical correlation for the discriminant function is .4883, which means that the variables in the function account for 23.8 percent of the total variance in whether the prosecution engaged in plea negotiations. Average story length is the strongest predictor of prosecutorial behavior, contributing more than a third of the variance accounted for by the function, 34.8 percent, which is 8.3 percent of the entire variance.

Easier to understand, perhaps, is the fact that the discriminant function correctly predicts the prosecutor's negotiating behavior in 69 of the 90 cases (76.7 percent). Without the information contributed by the variables in the function, successful predictions could be made only half of the time. In other words, knowing the values of the variables in the function provides a 53.3 percent improvement in predictive ability over chance guessing. The function is statistically significant at the .0003 level.

Despite the statistically significant results of the quantitative analysis, the cases in which the discriminant function makes an incorrect prediction merit some attention. In 5 of those 21 cases, the discriminant function computes a case's probability of membership in a given category (negotiated or not negotiated) at more than 80 percent.[2]

There was no explicit evidence of negotiations in two of these extreme outlier cases (*Wisconsin* v. *Reynosa, Wisconsin* v. *Shelton*), despite circumstances in both—minimal press attention, previous relationship with victim, no prior criminal record, etc.—that strongly predict bargaining. Both cases went to trial, and both defendants were convicted.

The striking fact about the cases, however, is that both defendants may sincerely have believed they were innocent. Virginia Reynosa, for example, did not actually stab the person whose death led to her first-degree murder conviction; she merely urged her son on as he wielded the knife. In addition, there was some evidence of provocation on the part of the victim. The son was convicted of murder for the stabbing; the mother was convicted under a portion of Wisconsin law that makes parties to a crime as guilty as the actual perpetrators. The charges are legally equivalent, however, and in Wisconsin a first-degree murder conviction carries a mandatory life prison sentence.

Jessie Shelton said that she had been the victim of domestic violence, and shot her husband in self-defense. The prosecutor who took her to trial for manslaughter and the jury that convicted her may have been skeptical of that claim; Jessie Shelton weighed 290 pounds, and her husband was blind. The judge, however, was more sympathetic. Saying he found an element of self-defense in the shooting, he placed Shelton on five years' probation rather than sending her to prison.

Incidentally, Shelton's husband was the most prominent defendant or victim in this study. He owned a taxi company serving Milwaukee's black community, and he had been active in community affairs. A measure of his prominence is that he was the only defendant or victim who had been the subject of non-crime-related news coverage before the homicide. Despite that—and despite Jessie Shelton's potentially newsworthy claim that she killed her husband in self-defense—pretrial press coverage of the case was minimal (a total of three relatively short routine stories.)

At least from a social point of view, Reynosa and Shelton may have believed, rightly or wrongly, that they were innocent. If so, it is not surprising they would be unlikely to be receptive to negotiations that would lead to guilty pleas and prison terms, as almost always is the case when a homicide case is plea bargained.

---

[2] It should be noted, of course, that in 19 other cases the discriminant function's 80 percent-or-better predictions were correct.

In the other three extreme cases (*Wisconsin* v. *Fraser*, *Wisconsin* v. *Crosley*, and *Wisconsin* v. *Murray*), prosecutors engaged in negotiations despite circumstances—intense press attention, extensive prior criminal record, multiple charges stemming from the incident, etc.—that strongly predict no negotiations. All three homicides were fairly routine. The press attention to the cases resulted from chance occurrences unrelated to the nature of the crime or of the people involved in it.

The Fraser and Crosley cases were fairly typical inner-city homicides; both began with holdups at all-night gas stations and ended with the shooting of the gas station attendants. In each case, the prosecution bargained with the defendant because it needed his testimony against an accomplice.

Those evidence-related considerations apparently outweighed the influence of extensive press coverage in the cases. But the press focused on the cases for reasons having little to do with the crimes themselves. Fraser's lawyer, for example, asked the judge to bar the press and public from the courtroom. The request, a relatively unusual one in Wisconsin, and the judge's denial of it inspired some press attention. The newspapers also singled out the Fraser case for exceptional feature coverage. The robbery and shooting took place around midnight Saturday, too late for more than a one-paragraph story in the final edition of Sunday morning's *Journal*. That gave reporters all day Sunday to interview nonroutine sources for substantial stories that would be published on Monday, typically a "slow" news day because governmental beats rarely generate news on Sunday.

In addition, five months after the incident, the *Journal* interviewed friends and family of the victim, and published a 51-paragraph Sunday feature titled "Last Day of Work—and Life." Before the suspects were arrested, the victim's family received a series of terrifying telephone calls. The *Journal* ran an 18-paragraph story ("Chilling Calls Taunted Victim's Kin").

In short, Fraser's crime was fairly routine; the coverage of it was not. It is a case that may have been plea bargained only because prosecutors had evidence problems that forced them to seek Fraser's agreement to testify against his accomplice.

The prosecution offered a similar bargain to Sylvania Crosley, charged with first-degree murder and armed robbery in connection with a crime similar to Fraser's. Crosley's crime took place in 1976, but no arrests were made until 1981, when an informant told police who had done it. By that time Crosley was in prison on unrelated charges; without his testimony, his alleged accomplice would go free.

Crosley initially agreed to the deal. But when it was time for him to testify, he refused and charges against his alleged accomplice in the gas-station holdup had to be dismissed. Defendants who renege on deals make prosecutors and judges unhappy. Crosley was convicted at trial of first-degree murder and armed robbery, and the judge imposed the maximum sentence—life plus 20 years.

Crosley's crime, like Fraser's, was fairly routine. The press coverage was routine, too, until allegations were made that money had been paid to Crosley's lawyer to buy Crosley's silence. The possibility of corruption sparked a 50-paragraph *Sentinel* story, "Probe of Payment to Lawyer Urged," and a 15-paragraph story a week later noting that authorities were probing ("Allegations on Payment of Lawyer's Fees Probed"). No other stories about the case were as many as 10 paragraphs long.

The final case that was negotiated despite intense press coverage stemmed from a tavern shooting that was routine in all but one respect—a group of Milwaukee *Sentinel* staffers was at the bar when the shooting started. Two of the journalists were seriously injured by stray bullets.

Because of weak and contradictory evidence, no arrests were made for more than a year after the shooting. Before any arrests were made, the *Sentinel* gave the case extensive publicity, including two lengthy first-person stories by one of the journalists who had been wounded. Other unusual stories during that period included a 63-paragraph story detailing victims' and witnesses' disappointment with the police department's failure to make arrests in the case and a 33-paragraph feature about the widow of the man who was killed in the shooting.

Fifteen months after the homicide, two suspects were arrested. The evidence was so weak that charges against one were dismissed outright. The other defendant, Regan Murray, faced first-degree murder and four nonhomicide charges, but agreed to plead guilty to reduced charges.

Why were the charges reduced when publicity had been so extensive? The available evidence gives no definite answer to that question, but it appears likely that the prosecution realized that the shaky evidence made conviction at trial problematic. The public pressure from the *Sentinel* made some kind of conviction imperative: thus the negotiations (Alschuler, 1968).

Discussion

This study's findings suggest that newspapers help set the plea bargaining agendas of Milwaukee prosecutors, at least in homicide cases. The amount of space newspapers were willing to devote to the typical story about a case was a stronger predictor than any other variable in this study of whether the prosecutor would negotiate.

This research found evidence suggesting an agenda-setting effect of ordinary crime news, the kind produced day in and day out by beat reporters. Recent studies from Northwestern University's Center for Urban Affairs and Policy Research, on the other hand, have examined possible agenda-setting effects only of investigative news reports (Cook et al., 1983; Protess et al., 1985).

In addition, this study did not use self-reported data. Most other research in this area has relied on public officials' own estimations of whether they have been influenced by the media (Gormley, 1975; Lambeth, 1978; Cook et al., 1983; Protess et al., 1985). Virtually all the measures of prosecutorial behavior used in this study, however, came directly from court records. The data came from cases that were already completed, ensuring that the participants were not aware that their behavior would become part of a research project. In addition, the data were recorded by court employees as part of the normal routine of the criminal courts, with no hint that the data would provide the raw material for an academic study. Unless the raw data contain a systematic bias that is not immediately apparent, they are unobtrusive measures (Webb, et al., 1966), unlike the self-reported data used as dependent measures in much of the research into the possibility that the news media may set the agenda for public officials.

That said, it must be acknowledged that documentary records are far from perfect. The facts of certain cases, such as the Reynosa and Shelton cases, strongly imply that the prosecution offered a deal that was turned down, despite the fact that court records and newspaper coverage contain no explicit evidence of negotiations.

Similarly, the examination of the five "outlier" cases revealed that the strength of evidence against a defendant (or a defendant's accomplice) could influence whether the prosecution negotiated a case. Evidence strength was not included as a quantitative variable in this study because it is extremely difficult to measure in any systematic way (Eisenstein and Jacob, 1977:182–183). In addition, evidence strength may well be a socially defined construct greatly dependent upon attributes of the relationships between the prosecutor and the defense lawyer, and between the defense lawyer and the defendant.

Finally, it is difficult to predict what effect a strength-of-evidence variable would have on plea bargaining. If evidence is weak, the prosecutor has an incentive not only to negotiate but to offer a good deal. If the evidence is strong, however, the defendant has an incentive to plead guilty even if not offered a good deal to avoid the so-called trial penalty (Uhlman and Walker, 1980; Shane-DuBow et al., 1981; Brereton and Casper, 1982; Pruitt and Wilson, 1983). Guilty pleas in cases where there is no evidence of negotiations often are implicit plea bargains (Heumann, 1978).

Related to the issue of evidence strength are questions involving the content of negotiations. Do prosecutors offer the same bargains to all similarly situated defendants, or does the defendant's race, how his or her case was covered, or other factors influence the content of the actual negotiations? This study did not have access to information about the offers and counter-offers that are a normal part of negotiations, but researchers should attempt to gain access to such information (see, e.g. Maynard, 1984).

Future research should also examine in detail the extent to which prosecutors (or other law-enforcement personnel) influence newspapers' decisions about which cases to cover and how to cover them (see, e.g., Drechsel, 1983). This study found no evidence of such an influence, but then court records and news coverage are not the best places to look for such an effect. Researchers need to get into newsrooms, prosecutors' offices, courthouse hallways, and courtrooms to find the answers to such questions.

This study focused on a narrow class of public officials—prosecutors handling homicide cases—in one community. As such, it offers no conclusions about whether the press sets the agenda for public officials in other situations and other places. Ideally, future studies will address such issues.

References

Alschuler, A.
1968 "The prosecutor's role in plea bargaining." University of Chicago Law Review 36:50–112.
1975 "The defense attorney's role in plea bargaining." Yale Law Journal 84:1179–1314.
Becker, L. B., M. E. McCombs, and J. M. McLeod
1975 "The development of political cognitions." In S. H. Chaffee, ed., Political Communication: Issues and Strategies for Research. Beverly Hills, Calif.: Sage Publications.
Blumberg, A.
1967 Criminal Justice. Chicago: Quadrangle Books.
Brereton, D. and J. D. Casper
1982 "Does it pay to plead guilty? Differential sentencing and the functioning of criminal courts." Law & Society Review 16:45–70.
Brosi, K. B.
1979 A Cross-City Comparison of Felony Case Processing. Washington, D.C.: U.S. Department of Justice.
Budd, R.
1964 "Attention score: a device for measuring news 'play'." Journalism Quarterly 41:259–262.

Buddenbaum, J. M., D. H. Weaver, R. L. Holsinger, and C. J. Brown
1981 Pretrial Publicity and Juries: A Review of Research. Bloomington, Ind.: Indiana University School of Journalism.
Bush, C. R.
1970 Free Press and Fair Trial: Some Dimensions of the Problem. Athens, Ga.: University of Georgia Press.
Casper, J.
1972 American Criminal Justice: The Defendant's Perspective. Englewood Cliffs, N.J.: Prentice-Hall.
Cohen, B. C.
1963 The Press and Foreign Policy. Princeton, N.J.: Princeton University Press.
Cohen, J., and P. Cohen
1975 Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. New York: John Wiley & Sons.
Cohen, S.
1975 "A comparison of crime coverage in Detroit and Atlanta newspapers." Journalism Quarterly 52:726–730.
Connors, M. M.
1975 "Prejudicial publicity: an assessment." Journalism Monographs 41.
Cook, F. L., et al.
1983 "Media and agenda setting: effects on the public, interest group leaders, policy makers, and policy." Public Opinion Quarterly 47:16–35.
Drechsel, R. E.
1983 News Making in the Trial Courts. New York: Longman.
Drechsel, R. E., K. Netteburg, and B. Aborisade
1980 "Community size and newspaper reporting of local courts." Journalism Quarterly 57:71–78.
Dunn, D.
1969 Public Officials and the Press. Reading, Mass.: Addison-Wesley.
Eisenstein, J., and H. Jacob
1977 Felony Justice: An Organizational Analysis of Criminal Courts. Boston: Little, Brown and Co.
Galanter, M.
1983 "Reading the landscape of disputes: what we know and don't know (and think we know) about our allegedly contentious and litigious society." UCLA Law Review 31:4–71.
Gifford, D. G.
1981 "Equal protection and the prosecutor's charging decision: enforcing an ideal." George Washington Law Review 49:659–719.
1983 "Meaningful reform of plea bargaining: control of prosecutorial discretion." University of Illinois Law Review 1983:37–98.
Gilberg, S., C. Eyal, M. McCombs, and D. Nicholas
1980 "The State of the Union Address and the press agenda." Journalism Quarterly 57:584–588.
Gormley, W. T., Jr.
1975 "Newspaper agendas and political elites." Journalism Quarterly 52:304–308.
Hearst Corporation
1983 The American Public, the Media & the Judicial System: A National Survey on Public Awareness and Personal Experience. New York: The Hearst Corporation.
Heumann, M.
1978 Plea Bargaining: The Experiences of Prosecutors, Judges, and Defense Attorneys. Chicago: University of Chicago Press.
Jones, J. B.
1978 "Prosecutors and the disposition of criminal cases: an analysis of plea bargaining rates." Journal of Criminal Law & Criminology 69:402–412.

Kaplan, J.
1965 "The prosecutorial discretion—a comment." Northwestern University Law Review 60:174–193.

Klecka, W. R.
1975 "Discriminant analysis." In Nie, N., et al. Statistical Package for the Social Sciences, second ed. New York: McGraw-Hill.

Lambeth, E. B.
1978 "Perceived influence of the press on energy policy making." Journalism Quarterly 55:11–18, 72.

Matthews, D. R.
1960 U.S. Senators and Their World. Chapel Hill: University of North Carolina Press.

Maynard, D. W.
1984 "The structure of discourse in misdemeanor plea bargaining." Law & Society Review 18:75–104.

McCombs, M., and D. L. Shaw
1972 "The agenda-setting function of mass media." Public Opinion Quarterly 36:176–187.

McLeod, J. M., L. B. Becker, and J. E. Byrnes
1974 "Another look at the agenda-setting function of the press." Communication Research 1:131–166.

Metropolitan Milwaukee Criminal Justice Council
1980 Public Opinion Survey 600.

Miller, F. W.
1970 Prosecution: The Decision to Charge a Suspect with a Crime. Boston: Little, Brown and Co.

Miller, S.
1978 "Reporters and congressmen: living in symbiosis." Journalism Monographs 53.

Milwaukee Journal
1981 "A deeper look at plea bargains." Editorial, January 11.

Newman, D. J.
1966 Conviction: The Determination of Guilt or Innocence Without Trial. Boston: Little, Brown and Co.

Peters, C.
1980 How Washington Really Works. Reading, Mass.: Addison-Wesley.

Pritchard, D.
1985 "Race, homicide and newspapers." Journalism Quarterly 62:500–7.

Protess, D. L., D. R. Leff, S. C. Brooks, and M. T. Gordon
1985 "Uncovering rape: the watchdog press and the limits of agenda setting." Public Opinion Quarterly 49:19–37.

Pruitt, C. R., and J. Q. Wilson
1983 "A longitudinal study of the effect of race on sentencing." Law & Society Review 17:613–635.

Rivers, W. L., S. Miller, and O. Grady
1975 "Government and the media." In S. H. Chaffee, ed., Political Communication: Issues and Strategies for Research. Beverly Hills, Calif.: Sage Publications.

Scott, W. A.
1955 "Reliability of content analysis: the case of nominal scale coding." Public Opinion Quarterly 19:321–325.

Shane-DuBow, S., et al.
1981 Wisconsin Felony Sentencing Guidelines: Phase I of Research and Development. Madison: Wisconsin Center for Public Policy.

Shaw, D. L., and M. E. McCombs
1977 The Emergence of American Political Issues: The Agenda-Setting Function of the Press. St. Paul: West Publishing.

Sherizen, S.
1978 "Social creation of crime news: all the news fitted to print." In C. Winick, ed., Deviance and Mass Media. Beverly Hills, Calif.: Sage Publications.
Sigal, L. V.
1973 Reporters and Officials: The Organization and Politics of Newsmaking. Lexington, Mass.: D. C. Heath.
Simon, R.
1977 "Does the court's decision in Nebraska Press Association fit the research evidence on the impact on jurors of news coverage?" Stanford Law Review 29:515–528.
Stanga, J. E.
1971 "The press and the criminal defendant: newsmen and criminal justice in three Wisconsin cities." Unpublished Ph.D. dissertation, University of Wisconsin.
Swank, D. H., H. Jacob, and J. Moran
1982 "Newspaper attentiveness to crime." In H. Jacob and R. L. Lineberry, eds., Governmental Responses to Crime: Crime on Urban Agendas. Washington, D.C.: National Institute of Justice.
Uhlman, T. M., and N. D. Walker
1980 " 'He takes some of my time, I take some of his': an analysis of sentencing patterns in jury cases." Law & Society Review 14:323–341.
Utz, P. J.
1976 Settling the Facts: Discretion and Negotiation in Criminal Court. Lexington, Mass.: Lexington Books.
Walker, J.
1977 "Setting the agenda in the U.S. Senate: a theory of problem selection." British Journal of Political Science 7:423–445.
Webb, E. J., D. T. Campbell, R. D. Schwartz, and L. Sechrest
1966 Unobtrusive Measures: Nonreactive Research in the Social Sciences. Chicago: Rand McNally & Co.
Weiss, C. H.
1974 "What America's leaders read." Public Opinion Quarterly 38:1–22.
Wills, R. H.
1977 Milwaukee Sentinel Editorial Policy.

Interpersonal Communication and News Comprehension

JOHN P. ROBINSON AND MARK R. LEVY

THE PERENNIAL survey finding that the American public obtains most of its information from television is widely cited and generally accepted. Indeed, almost two-thirds of all Americans now identify television as their "main source" of news and that proportion has grown steadily since the late 1950s (Roper, 1983, 1985). While numerous researchers have raised methodological cautions about this generalization (Carter and Greenberg, 1963; Stone, 1969–70; Robinson, 1971; Stevenson and White, 1980), the "debate" over where the public gets most of its news by and large has been conducted in a conceptual and theoretical vacuum.

First of all, "main source" studies are almost always based on self-reports of information gain. Such subjective measures indicate only where people *think* they find out the news. However, self-reports of

---

Abstract This study examines the role of interpersonal and mass media channels in public awareness and comprehension of major news stories. Unlike previous research which asks about respondent perceptions of their "main source" for news, this study attempts to determine *actual* comprehension of stories that had been in the news during the previous week. On the basis of two separate probability samples, respondent awareness and comprehension of a week's news was measured and related to demographic, news media use, and interpersonal discussions variables. A Multiple Classification Analysis of the data indicates that conversation about the news is a major and often overlooked correlate of comprehension, and that interpersonal channels may play at least as important a role in the public's awareness and understanding of the news as exposure to the news media.

John P. Robinson is the Director of the Survey Research Center and a Professor of Sociology at the University of Maryland. Mark R. Levy is an Associate Professor of Journalism and a Research Associate of the Center for Research in Public Communication, University of Maryland. An earlier version of this article was presented to the 40th annual meeting of the American Association for Public Opinion Research, McAfee, NJ, May 1983. Support for this study was provided by the John and Mary Marke Foundation.

Public Opinion Quarterly Vol. 50:160–175 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-160/$2.50

channel influence measure neither *actual* information gain nor, perhaps more important, how well that information is comprehended. These shortcomings have hindered our ability to understand the role of the mass media in the opinion formation process and suggest the need for studies based on better measures of information gain and comprehension.

Second, when asked "main source" questions, fewer than 1 respondent in 20 typically points to interpersonal channels as their principal news source. This finding stands in marked contrast to several well-established paradigms of public opinion and mass communication research. For example, the importance of "people" as an information source has been demonstrated in numerous studies of political communication effects (e.g., Katz and Lazarsfeld, 1955; Tichenor et al., 1973; Robinson, 1976), diffusion of innovations (e.g, Mason, 1963; Rogers, 1983), and communication about "major" news stories (e.g., Troldahl and Van Dam, 1965; Greenberg, 1964; Gantz, 1983). On the basis of this literature, then, one might expect to find previously overlooked linkages between interpersonal communication, information gain, and news comprehension.

The study presented here examines the comprehension of news and relates levels of comprehension to the use of mass media and interpersonal communication channels by individuals. This study is informed by the revision of the two-step flow hypothesis outlined in Robinson (1976) and by the general literature on interpersonal communication (see, for example, Schramm, 1973; Chaffee, 1982). It assumes that information is often, but not always, characterized by a "horizontal" flow between reasonably well-informed and interested individuals, and that information exchange in interpersonal settings is facilitated by such factors as greater personal involvement, immediate feedback, tailoring of messages, and the like.

We further assume that interpersonal discussions may be an important and largely overlooked agent in what Gerbner et al. (1980) have hypothesized is a "mainstreaming" process. Given that much of what individuals know about "the world out there" comes to them from the mass media *and* through interpersonal channels, the interaction of media messages and interpersonal discussion becomes central in considering the mechanisms that create possible "mainstreaming" effects. It is, after all, well documented that interpersonal discussions sometimes exercise a social control function, pressuring individuals into conformity with generally accepted perceptions and understandings. If those attitudes, perceptions, and understandings originate even partly in media messages, then talking about the news may simultaneously produce audience understandings which are widely shared and which are highly similar to journalistically created accounts of social reality.

Method

Unlike previous studies of news comprehension which have experimentally manipulated message attributes (e.g., Findahl and Höijer, 1976; Katz et al., 1977; Gunter, 1980) or which have sought to determine recall and comprehension of a single evening's news (e.g., Neuman, 1976; Robinson et al., 1982), this study focuses on actual national and international news messages, published and broadcast over a week's time.

While there is nothing magical about a seven-day period, asking about news at the end of a week may more realistically reflect actual patterns of audience exposure. The chance, for instance, that an individual will have watched any given evening's network news may be as small as 1 in 20 (Robinson, 1971; Lichty, 1982). However, over the course of a week, the likelihood of any exposure to network news and its associated opportunity for information gain and comprehension rises substantially.

Additionally, over the course of a week, there is an accumulation of exposure to the news (Bogart, 1981). To the extent that coverage of "important" news within and across news channels is complementary or redundant, audiences are more likely to receive several accounts of the same news item—and that too might aid awareness and comprehension. Further, taking a week's worth of news virtually assures that a greater substantive variety of news will have time to unfold and develop. Increased story "maturity" may aid individuals seeking to integrate the news into semantic memory (Woodall et al., 1983). Finally, studying what people know about the news after a week may allow time for interpersonal discussion of the news. News stories sometimes serve as a "coin of social exchange" (Levy, 1978a, 1978b), and, as was suggested above, it is possible that the interpersonal utility of the news may enhance its recall and comprehension.

SAMPLES

Two separate probability surveys were conducted by telephone from the Survey Research Center at the University of Maryland using the random-digit-dialing method of household selection. In the first sample, 407 adults living in the greater Washington, DC metropolitan area were interviewed during May 1983. To maximize generalizability and to obtain a greater variety of news stories, a second set of interviews with 544 adults nationwide were carried out during June of 1983.

Most interviewing was done on weekends and was based on the stories that had received prominent media attention in the preceding seven days. Interviews were spread approximately equally throughout each month. Completion rates were 74 percent for the regional sample and 71 percent for the national sample, after 10 callbacks to establish whether the telephone numbers were actually of telephone households.[1] Data from both surveys were weighted by respondent sex, age, and education to ensure representativeness with 1980 Census Bureau figures for the region and the nation.

NEWS MEDIA USE QUESTIONS

In both surveys, respondents were asked the following media exposure questions: (1) "Over the past seven days, have you read any daily newspaper?" (2) "Over the past seven days, have you read any national news or commentary magazine?" (3) "Over the past seven days, have you listened to any news broadcasts on the radio?" (4) "Over the past seven days, have you watched any news programs on television?" Each "yes" response was coded "1"; each "no" was coded "0." Responses on these four items were summed for each respondent into an index of news exposure.

All respondents were also asked, "Did you talk with anyone about anything in the news last week?" If "yes," respondents were then asked, "About how many such conversations did you have?" and "Who did you talk with?" Additionally, respondents reporting conversations about the news were asked: "Can you remember any of the things you talked about?" and if so, "What was that?"

News stories discussed by respondents in the national sample were then coded by two trained research assistants and one of the authors to determine whether the item was "definitely," "maybe," or "not" one of the "most important" national or international stories of the preceding week. (A discussion of how the week's "most important" stories were identified follows below). Intercoder agreement on this measure was 87 percent.

While respondents were free to mention conversations about local, national, or international news items, our analysis focused on "cosmopolitan" news (Merton, 1949), that is, national and international stories of more than purely local interest. While knowledge of different types of news may have different correlates, and awareness and comprehension of local news may be even more influenced by word of mouth, at least one study that systematically examined *both* national and local news stories has found considerable similarity in the background and media use correlates associated with both types of news (Becker and Whitney, 1980).

---

[1] Estimates of public comprehension of the news presented below are, of course, subject to sample biases due to the elimination of nontelephone households and the nonresponse rate. Telephone interviewing may have excluded, for example, some poorer, older, and non-English-speaking respondents, while some people who are less interested in the news and less well informed about it may have disproportionately refused to be interviewed. Thus, although it should not affect our analysis of the *correlates* of news comprehension, it is likely that our point estimates of news comprehension may be marginally inflated, even though the data were weighted to match census figures for education, age, and sex distribution.

In addition, to assess and control for individual interest in the news, respondents in the national sample were asked, "How closely do you follow national and international news in the media? Would you say very closely, somewhat closely, or not very closely." Respondents were also asked standard demographic items (e.g., sex, age, education, etc.).

SELECTION OF NEWS COMPREHENSION ITEMS

All respondents were asked "awareness" and "comprehension" questions about the "most important" stories from the week preceding the interview. The selection of "most important" news stories represented a collective judgment by a rotating panel of five or more high-ranking news professionals who worked for *The Washington Post*, *USA Today*, Associated Press, CBS News, and ABC News. (For a more general discussion of this technique, see Robinson et al., 1982.) Each week this panel of journalists was asked: (1) What have been the "big stories" of the past week? (2) What were the one or two most important things ("main points") that the audience should have learned from those stories?

Examples of "important stories" and their "main points" include: an earthquake in California with no fatalities; the nationwide outbreak of AIDS and its effect on homosexuals, Haitians, and others; Congress limiting the "tax cut" out of concern over the federal deficit; U.S. Senator John Glenn announcing his presidential candidacy; the Pope visiting Poland to support Solidarity; the reelection of British Prime Minister Margaret Thatcher by a landslide; and the expulsion of the U.S. ambassador from Nicaragua for "spying" and the subsequent reaction of the Reagan administration.[2]

In choosing the final list of stories for inclusion in the week's survey, every effort was made to include only items that had received extensive coverage in both print and broadcast news media. Finally, the wording of survey questions was checked with the panel to insure appropriateness and accuracy. Audience awareness of any given news story was measured by a "yes–no" response to a question which took the generic form: "Did you hear or read anything last week about ——?" If "yes," the respondent was then asked a question, usually open-ended, which had been designed to test the degree to which the individual knew a given story's "main point" as defined by the panel of journalists.[3] Over the course of each four-week field period, we asked respondents in the regional sample about 23 different persons, events, or phenomena that had been in the news, and respondents in the national sample about 31 distinct persons, events, or phenomena.

---

[2] For a complete listing of news stories and their associated awareness and comprehension questions, see Robinson and Levy (1986).

SCORING OF NEWS COMPREHENSION ITEMS

Responses to the open-ended comprehension items were coded by the authors using a 9-point scale, developed in connection with earlier comprehension studies (Robinson and Sahin, 1984). Coding was by consensus, with initial intercoder agreement in excess of 90 percent. Scores on the comprehension scale ranged from 0 to 8. Responses which recalled either the "wrong" story or failed to give its "main point" were rated 0 or 1, while a failure to recall anything beyond the initial awareness of the story was coded 2. Vague recollections of the story were coded 3 or 4, depending on response specificity. The code of 4 was assigned if the respondent was able to recall correctly some detail about the story (e.g., the identity of a the principal person in the story), but not the story's "main point." A reply containing the correct "main point," however, was coded 5, with one point added to the basic score of 5 for any additional correct detail provided—up to a score of 8 if the respondent noted not only the "main point" but three additional, cogent story details.

Thus, for example, in stories about AIDS, respondents were asked, "Do you know about any particular group of people who have been affected (by AIDS) more the others?" A response of "homosexuals" was coded 5. If the respondent said "Homosexuals" and Haitians, drug addicts, and/or hemophiliacs, that response was scored as a 6, 7, or 8, depending on how many groups beyond the "main point" (homosexuals) were named.

A summary measure representing total news comprehension for the week was then constructed for each respondent.[4] For each answer to a specific comprehension item scored 5, 6, 7 or 8, the respondent received one point. These points were then summed and normalized to equalize the total number of weekly news items which varied slightly from week to week. Scores on this summary measure of news comprehension ranged from 0 to 9 in the regional sample and from 0 to 14 in the national sample. A score of 0 indicated no stories comprehended, while a summary score of 9 in the regional sample or 14 in the national study indicated comprehension of all items for a given week.

---

[3] Examples of "main point" comprehension questions include: "How many people were killed in the (California) earthquake?" "Why did Congress want to limit the tax cut?" "Why was Senator Glenn in the news?" "What do you think the purpose of the Pope's trip (to Poland) was?

[4] Obviously, two comprehension measures have been developed: the first gauges the "depth" of information on a single story; the second, the summary score, measures the

Ideally for our research, we should have had a specially devised "information meter" that would have automatically registered when respondents understood a news story and that recorded the source from which the information came. Given the current unavailability of such a device, we simulated the process with surveys. Our simulation asked first about the news information that respondents did have and second about the news sources respondents used. If users of news medium A possessed information that users of news media B did not, we took that to be reasonable evidence that news medium A was the more important information source. While this is not causal evidence (if indeed such social science data can be definitively gathered at all), it is the closest *field* approximation that we believed could be devised.

Results

The mean "awareness" score was 68 percent in the regional sample and 62 percent in the national sample. That means that roughly two-thirds of respondents in each sample said they had heard or read about the story. Details of U.S. government actions, in both domestic and foreign policy, tended to rank near the bottom in terms of public awareness, while respondents were most aware of especially dramatic news of real or potential danger and "human interest" stories.[5]

Means news comprehension scores were 2.7 (out of 9 weekly stories) in the regional sample, and 4.6 (out of an average of 14 stories per week) in the national sample. Thus, on average, somewhat less than one-third of the major news stories examined each week were comprehended by respondents sampled. No respondents in the regional sample successfully understood the "main point" of more than five of the week's major news stories. Fewer than 4 percent of national re-

---

"breadth" of respondent comprehension *across* a number of news stories. Limited analysis to date has found that the media use and interpersonal communication correlates of the two comprehension measures are highly similar.

[5] A story-by-story listing of "awareness" and "comprehension" scores is available on request from the authors.

Table 1. Weekly Conversations About News

	Regional Sample (407)	*National Sample (544)*
No. of conversations		
None	35%	40%
1–2	10	14
3–5	25	23
6–9	16	12
10 or more	14	11
Discuss "big" news		
Yes	n.a.	18
Maybe	n.a.	15
No	n.a.	51
Can't recall topic(s)	20	16

spondents comprehended 12 or more items and none had a score of 14. In general, these results are consistent with earlier parallel studies conducted in Great Britain (Robinson and Sahin, 1984); and like the pattern reported above for the awareness measure, comprehension was often, but not exclusively, higher for stories with dramatic or "human interest" content.

Almost two-thirds (65 percent) of respondents in the regional sample and almost as many (60 percent) of national respondents said they had talked "about things in the news" during the week prior to their being interviewed (Table 1). In both samples, the largest proportion of conversations, two-thirds or more, were with "friends" or "colleagues at work," with family members making up the remainder of conversational partners.

Only 20 percent of regional respondents and 16 percent of the national sample who reported having news-oriented conversations were unable to recall the topic of those discussions. This is a rather low proportion, based as it is on an unaided recall question. Indeed, it suggests that news—be it local, national, international, or some mix—has a relatively high salience for most individuals.

Of respondents (national sample only) who could recall what they talked about, some 21 percent reported having at least one conversation which was coded as "definitely" dealing with one of the "most important" stories of the week. Another 18 percent of the national sample reported at least one conversation that was coded as "maybe" focusing on one of the week's major stories.

BIVARIATE CORRELATES OF NEWS COMPREHENSION

The average scores on the summary measure of news comprehension for each major background variable group are presented in Table

Table 2. Mean News Comprehension Scores by Demographic, Media Use, and Communication Behavior Variables

	Regional Sample		*National Sample*	
All respondents	2.7	(407)	4.6	(544)
Sex				
Male	3.3	(180)	5.7	(237)
Female	2.3	(226)	3.8	(307)
Age				
18–19	0.7	(9)	2.9	(17)
20–29	2.3	(97)	3.8	(127)
30–39	3.2	(90)	4.8	(113)
40–49	3.2	(63)	5.1	(82)
50–59	2.7	(68)	5.0	(63)
60–69	2.6	(49)	5.4	(76)
70 or older	2.7	(31)	4.2	(66)
Education				
< High school	1.7	(80)	3.1	(95)
H.S. grad	2.1	(119)	4.0	(278)
Some college	2.7	(71)	5.4	(94)
College grad	4.0	(137)	6.7	(49)
Graduate educ	—[a]		8.4	(28)
News exposure index				
0	0.5	(6)	1.2	(3)
1	1.0	(37)	2.8	(51)
2	2.1	(76)	3.6	(156)
3	2.8	(177)	5.0	(234)
4	3.6	(110)	6.2	(100)
How closely follow news				
Very closely	n.a.		6.4	(101)
Somewhat	n.a.		5.3	(266)
Not very	n.a.		2.5	(177)
Talks about news				
None	1.8	(143)	3.3	(219)
1–2	2.3	(38)	4.2	(74)
3–5	2.8	(101)	5.5	(127)
6–9	3.8	(66)	5.7	(64)
10 or more	3.9	(59)	6.8	(59)
Discuss "big" news				
Yes	n.a.		6.8	(98)
Maybe/no/DK	n.a.		4.1	(446)

[a] Respondents with post-college education included in college graduate category.

2. In general, there is considerable consistency between the regional and national samples with regard to these correlates of news comprehension. In both surveys, for example, there were fairly large differences in the comprehension scores of men and women, with men tending to have higher scores. Comprehension also increased with respondent age, although both sets of data showed some tendency for curvilinearity, with older respondents scoring lower than middle-aged respondents.

In both samples, level of formal education was the major predictor of news comprehension (see Robinson, 1967; Hyman et al., 1977 for similar results with other information survey questions). The average comprehension score of respondents with a college degree, for example, was more than double that of respondents who had graduated from high school.

Exposure to news media was also a clear but somewhat less powerful correlate of comprehension scores. For example, respondents in both samples with the least exposure to the news (i.e., 0 or 1 on the News Exposure Index) had mean scores that were less than half those achieved by those respondents reporting exposure to all four media news sources. However, because of the small size of this minimal exposure group, news exposure had less of a correlation with comprehension scores than did formal education. Interest in the news was another important correlate of comprehension scores, with respondents who claimed to follow the news "very closely" also scoring over twice as high as those who followed the news "not very closely."

Average news comprehension scores were also clearly correlated with the interpersonal communication variable. Talking about news with other people was related to considerably higher comprehension levels. As the number of conversations about news increased from none to 10 or more for the week, the average comprehension score more than doubled in both samples. Additionally, respondents in the national sample who reported talking about one or more major stories of the week had comprehension scores which were among the very highest in the study.

Table 2, then, links news comprehension scores with several possible predictor variables, considering each of those variables one at a time. The question naturally arises whether some of these correlates emerge as more powerful predictors after simultaneous adjustments are made for each of the other correlates in Table 2. For example, many of the correlates of comprehension examined here are also often linked with the strongest correlate of comprehension, namely, formal education. It could well be that controlling for the respondent's level of education would substantially reduce, or even eliminate, the correlations of news comprehension with news media exposure, interest in the news, or news-oriented conversations.

MULTIVARIATE ANALYSIS OF NEWS COMPREHENSION

News comprehension scores were next subjected to a Multiple Classification Analysis (Andrews et al., 1973) in order to determine the impact on comprehension of media use and communication behavior variables after controlling for demographic factors and the other media use questions. This MCA technique provides a combination multiple regression and analysis of variance adjustment on survey data.[6]

As Table 3 demonstrates, differences in the average comprehension scores for men and women in both samples are substantially reduced, after the other factors in Table 2 are taken into account. Statistical controls also reduce some of the curvilinearity of news comprehension and age in the regional sample, while educational differences and comprehension in both samples are also somewhat reduced.

Differences by levels of news media exposure are also considerably constricted by multivariate controls; and similarly, the gap is narrowed between respondents in the national sample who follow the news "very closely" and only "somewhat closely."

The interpersonal discussion variable is also affected by simultaneous controls for other variables, and several departures from the earlier pattern are found as well. Thus, in the regional sample, peak comprehension scores are reached in the 6–9 conversation group and then dip slightly in the 10 or more conversation category. In the national sample, peak comprehension scores are reached in the 10+ conversation group, while the mean comprehension score for respondents reporting 6–9 conversations dips below the score for the 3–5 conversation group.

The small sample sizes for the two groups reporting the largest number of conversations may account for some of the instability in the data. Indeed, if the data from the two samples in Table 3 were merged, the combined data would show clear monotonicity with number of conversations. The difference in average comprehension scores between the no-conversation group and the 10+ conversation group then would be 1.1 for the regional sample and 1.6 for the national sample, representing more than a 30 percent gain in news comprehension after MCA adjustments. For the national sample that is a larger difference than the 0.8 point gain between the four-media vs. 0–1 media use group combined (to obtain meaningful sample size); for the regional sample, the 1.4 point gain in mean comprehension scores across news media use is 0.3 points larger, but that is not a large difference.

---

[6] One advantage of MCA over other multiple regression techniques is that it provides straightforward adjusted averages for specific categorical groups in the sample (e.g., for men and women, or blacks, whites and other racial groups.) An MCA is also useful for ordered groups (e.g., high vs. middle vs. low education), since one can see whether relationships are linear and monotonic or have "peaks" and "valleys" in the middle categories. The principal drawback of MCA is that it does not provide error estimates that are necessary to conduct standard statistical tests on average difference scores between groups. However, certain "rules of thumb" have been developed for interpreting average group differences, and they are followed, where applicable, in this article.

Table 3. MCA Analysis of Comprehension Scores by Demographic, Media Use, and Communication Behavior Variables

	Regional Sample	*National Sample*
All respondents	2.7 (407)	4.6 (544)
Sex		
Male	3.0 (180)	5.3 (237)
Female	2.6 (226)	4.0 (307)
Age		
18–19	1.9 (9)	3.9 (17)
20–29	2.3 (97)	3.8 (127)
30–39	2.9 (90)	4.3 (113)
40–49	2.7 (63)	4.9 (82)
50–59	2.8 (68)	5.0 (63)
60–69	2.9 (49)	5.6 (76)
70 or older	3.7 (31)	4.8 (66)
Education		
< High school	2.0 (80)	3.4 (95)
H.S. grad	2.2 (119)	4.3 (278)
Some college	2.7 (71)	5.2 (94)
College grad	3.7 (137)	6.1 (49)
Graduate educ	—[a]	7.1 (28)
News exposure index		
0	2.0 (6)	3.3 (3)
1	1.6 (37)	4.1 (51)
2	2.4 (76)	4.3 (156)
3	2.9 (177)	4.8 (234)
4	3.1 (110)	4.8 (100)
How closely follow news		
Very closely	n.a.	5.6 (101)
Somewhat	n.a.	3.0 (266)
Not very	n.a.	3.5 (177)
Talks about news		
None	2.3 (143)	4.1 (219)
1–2	2.3 (38)	4.2 (74)
3–5	2.6 (101)	5.1 (127)
6–9	3.6 (66)	4.9 (64)
10 or more	3.4 (59)	5.7 (59)
Discuss "big" news		
Yes	n.a.	5.7 (98)
No	n.a.	4.4 (446)

[a] Respondents with post-college education included in college graduate category.

Thus, in both samples, the extent of discussion of the news seems to be at least as powerful a predictor of comprehension as the extent of news media exposure, and in the more generalizable national sample, such discussion was associated with almost twice as much spread in news comprehension as was media exposure.

A final insight into the importance of the interpersonal factor is given at the bottom of Table 3. If a respondent reported talking about even one of the week's "big news" stories, then, even after controlling for all other correlates reported in the table, that respondent was likely to score significantly higher on comprehension. Coupled with the findings reported above about comprehension and the number of news-oriented discussions, this result suggests that both the quantity *and* the quality of conversations are important factors in understanding the news.

Discussion

In summary, then, what can be said about the public's awareness and comprehension of national and international news over a week's time? First, and perhaps most obviously, it should be noted that there is a striking similarity between findings from the two surveys. Despite considerable geographic differences among respondents in the two samples and despite differences in the news stories examined each week, there is considerable overlap in the patterns within and the relative importance between key news comprehension correlates. This consistency across two independent samples points to a reassuring degree of generalizability in our findings.

Second, while education was found to be the single most important predictor of news comprehension, respondents who reported more conversations about the news were also more likely to score higher in news comprehension. Talking about the news, especially the week's "most important" national and international news, seems to have at least as powerful an effect on comprehension as mere exposure to the news media.

That finding, we believe, has serious implications for journalism, public understanding, and even the democratic process. While a full discussion of those implications is well beyond the scope of this article, we want to make one observation in passing. Journalists and politicians often complain that the public is not interested in the news and lacks the educational background to understand what gets reported (Robinson et al., 1982). Our study demonstrates, however, that controlling for news interest and education, public understanding of the news does increase as people talk about it. Journalists and others would do well, then, to consider ways to create and present news which stimulates or otherwise takes advantage of this powerful interpersonal "second stage" of the information flow.

Finally, the findings presented here bear directly on our general theories and models of the mass communication process. Two general paradigms inform most communication scholarship: one, the so-called dominant or American approach, is a fundamentally empirical and pluralistic model of relatively limited communication effects in which the power of the mass media is often mitigated by interpersonal influences; the other, the so-called critical or European perspective, finds major conceptual and empirical flaws in the dominant model and assumes often pervasive power for the media (Curran et al., 1982; Gitlin, 1978).

Our study, we believe, suggests a bridge between the two paradigms. On the one hand, our findings strongly suggest that interpersonal communication plays a crucial role in the reception, retransmission, and interpretation of mass media messages. That conclusion argues in support of the dominant theory and its continuing emphasis on the importance of social networks for the mass media process.

However, that same conclusion, namely that interpersonal communication increases the fidelity between the manifest content of mass media messages and the interpretations and meanings derived from those messages by audiences, can be interpreted to support the notion of a more powerful, "mainstreaming," or possibly even "ideological effect" of mass communication (Hall, 1983). From this perspective, interpersonal communication could be seen as helping to create audience "readings" of mass media messages that are closer in substance to "professional," and thereby ideologically "preferred," understandings of reality (Hall, 1980).

To substantiate such an assertion, of course, would require findings drawn from many different research sites in the mass communication process, and would necessarily involve both micro- and macro-studies of audiences, communicators, and culture more generally. This article obviously has a more limited scope, although our findings do have implications for such an investigation.

We have not, for example, offered *direct* evidence that the public interprets the news from any particular ideological standpoint; and we are aware that audiences are at least capable of creating "negotiated" or "oppositional" decodings of news messages (Morley, 1980). But we also have shown that public awareness and comprehension of the news, to the extent it occurs at all, often bears a striking similarity to journalistically encoded reality.

Similarly, the complex and often problematic relationship between journalistic encodings of the news and the more general phenomenon of ideology remains unexamined here. But other researchers have already shown that the ideological and thematic range of most news messages is fairly limited (Gans, 1979), and that represents an important precondition for demonstrating a linkage between audiences, communicators, and ideologies.

Taking these factors into account, and fully recognizing that much more work is needed, we still think it is plausible to speculate that conversations about the news will both increase the "accuracy" with which news messages are comprehended and the degree to which those messages may have ideological, essentially status quo consequences.

References

Andrews, F. M., J. N. Morgan, J. A. Sonquist, and L. Klem
1973 Multiple Classification Analysis. Second Edition. Ann Arbor: Institute for Social Research, The University of Michigan.
Becker, Lee, and D. C. Whitney
1980 "Effects of media dependencies." Communication Research 7:95–120.
Bogart, L.
1981 Press and Public. Hillsdale, NJ: Lawrence Erlbaum Associates.
Carter, R., and B. Greenberg
1965 "Newspapers or television: which do you believe?" Journalism Quarterly 42:29–34.
Chaffee, S.
1982 "Mass media and interpersonal channels: competitive, convergent, or complementary?" In G. Gumpert and R. Cathcart (eds.), Inter/media. Second Edition. New York: Oxford University Press.
Curran, J., M. Gurevitch, and J. Woollacott
1982 "The study of the media: theoretical approaches." In M. Gurevitch, T. Bennett, J. Curran, and J. Woollacot (eds.), Culture, Society and the Media. New York: Methuen.
Findahl, O., and B. Höijer
1976 Fragments of Reality: An Experiment with News and TV Visuals. Stockholm: Sveriges Radio.
Gans, H.
1979 Deciding What's News. New York: Pantheon.
Gantz, W.
1983 "The diffusion of news about the attempted Reagan assassination." Journal of Communication 33 (Winter):56–66.
Gerbner, G., L. Gross, M. Morgan, and N. Signorielli
1980 "The 'mainstreaming' of America: Violence Profile No. 11." Journal of Communication 30:10–29.
Gitlin, T.
1978 "Media sociology: the dominant paradigm." Theory and Society 6:205–253.
Greenberg, B.
1964 "Person to person communication in the diffusion of news events." Journalism Quarterly 41:489–494.
Gunter, B.
1980 "Remembering the television news: effects of visual format on information gain." Journal of Educational Television 6:8–11.
Hall, S.
1980 "Encoding and decoding." In S. Hall, D. Hobson, A. Lowe, and P. Willis (eds.), Culture, Media, Language. London: Hutchinson.
1983 "The rediscovery of 'ideology': return of the repressed in media studies." In M. Gurevitch, et al. (eds.), Culture, Society and the Media. New York: Methuen.
Hyman, H., C. Wright, and J. Reed
1977 The Enduring Effects of Education. Chicago: University of Chicago Press.
Katz, E., H. Adoni, and P. Parness
1977 "Remembering the news: what the picture adds to recall." Journalism Quarterly 54:231–239.
Katz, E., and P. Lazarsfeld
1955 Personal Influence. New York: The Free Press.
Levy, M.
1978a The Audience Experience with Television News. Association for Education in Journalism and Mass Communication, Journalism Monographs, Number 55.
1978b "Opinion leadership and television news uses." Public Opinion Quarterly 42:402–406.

Lichty, L.
1982 "Video vs. print." The Wilson Quarterly, Special Issue.
Mason, R.
1963 "The use of information sources by influentials in the adoption process." Public Opinion Quarterly 27:455–466.
Merton, R.
1949 "Patterns of influence: a study of interpersonal influence and communication in a local community." In P. Lazarsfeld and F. Stanton (eds.), Communication Research 1948–1949. New York: Harper and Brothers.
Morley, D.
1980 The 'Nationwide' Audience. London: British Film Institute.
Neuman, W. R.
1976 "Patterns of recall among television viewers." Public Opinion Quarterly 40:115–23.
Robinson, J.
1967 "World affairs information and mass media exposure." Journalism Quarterly 44:23–30.
1971 "The audience for national TV news programs." Public Opinion Quarterly 35:403–05.
1976 "Interpersonal influence in election campaigns: two-step flow hypotheses." Public Opinion Quarterly 40:304–19.
Robinson, J., and M. Levy
1986 The Main Source: Learning from Television News. Beverly Hills, CA: Sage Publications.
Robinson, J., and H. Sahin
1984 Audience Comprehension of Television News: Results from Some Exploratory Research. London: BBC Audience Research Department.
Robinson, J., H. Sahin, and D. Davis
1982 "Television journalists and their audiences." In J. Ettema and D.C. Whitney (eds.), Individuals in Mass Media Organizations. Beverly Hills: Sage Publications.
Rogers, E.,
1983 Diffusion of Innovations. Third Edition. New York: Free Press.
The Roper Organization
1983 Trends in Attitudes Toward Television and Other Media: A Twenty-Four Year Review. New York: Television Information Office.
1985 Public Attitudes Toward Television and Other Media—A Time of Change. New York: Television Information Office.
Schramm, W.
1973 "Channels and audiences." In I. de Sola Pool and W. Schramm (eds.), Handbook of Communication. Chicago: Rand McNally.
Stevenson, R., and K. White
1980 "The cumulative audience of network television news." Journalism Quarterly 57:477–81.
Stone, V.
1969–1970 "Sources of most news: evidence and influence." Journal of Broadcasting 14:1–4.
Tichenor, P., J. Rodenkirchen, C. Olien, and G. Donohue
1973 "Community issues, conflict, and public affairs knowledge." In P. Clarke (ed.), New Models for Communication Research. Beverly Hills, CA: Sage Publications.
Troldahl, V., and R. Van Dam
1965 "Face to face communication about major topics in the news." Public Opinion Quarterly 29:626–34.
Woodall, W., D. Davis, and H. Sahin
1983 "From the boob tube to the black box: television news comprehension from an information processing perspective." Journal of Broadcasting 27 (Winter):1–23.

Measuring Change in Personal Economic Well-Being

STEVEN J. ROSENSTONE, JOHN MARK HANSEN
AND DONALD R. KINDER

FOR OVER A decade, researchers have tried to unravel the political consequences of economic conditions—with mixed success. Studies using aggregate, over-time, data have found that prosperity enhances the electoral chances of incumbents and bolsters support of chief executives.[1] Parallel investigations at the individual level, however, have turned up a striking anomaly: changes in citizens' personal economic well-being seem to have little or no bearing on their political preferences.[2]

Efforts to reconcile these findings have ventured down several paths. Perhaps citizens respond not to changes in their own personal economic situation but to evaluations of the economic health of the nation as a whole (Kinder and Kiewiet, 1979, 1981) or the economic

---

Abstract Individual-level investigations of the link between personal economic well-being and political preference commonly turn up findings at odds with aggregate, over-time analyses. Although various theoretical explanations for this anomaly have been advanced, we focus on a different culprit: the possibly frail measure employed in survey based studies. We develop improved measures of change in personal economic well-being and use a national survey of the American electorate to test them. Our findings are striking: although the traditional question used by survey researchers is quite valid, its unreliability has caused the effect of personal economic well-being on political evaluations to be substantially underestimated.

Steven J. Rosenstone is Professor of Political Science at Yale University. John Mark Hansen is Assistant Professor of Political Science at The University of Chicago and is a Research Fellow at The Brookings Institution. Donald R. Kinder is Associate Professor in the Department of Political Science and Associate Research Scientist at the Institute for Social Research, both at the University of Michigan.

[1] E.g., Kramer, 1971; Bloom and Price, 1975; Frey and Schneider, 1975; Tufte, 1978; Kernell, 1978; Hibbs and Fassbinder, 1981; Hibbs, 1982; and Rosenstone, 1983.

[2] E.g., Fiorina, 1978, 1981; Kinder and Kiewiet, 1979, 1981; Schlozman and Verba, 1979; Sears et al., 1980; Kinder, 1981; Kiewiet, 1983.

Public Opinion Quarterly Vol. 50:176–192 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press
0033-362X/86/0050-176/$2.50

predicaments of groups (Conover, 1985; Kinder et al., 1983). Perhaps personal finances have political consequences, but only for rather unusual voters (Feldman, 1982; Kinder and Mebane, 1983). Or perhaps the political effect of family finances has to do not with the voter's calculus but with the caliber of candidates: the appearance of attractive, well-financed out-party candidates hoping to reap windfalls from the incumbents' economic woes (Fiorina, 1978; Jacobson and Kernell, 1981).[3]

We take a different route. We argue here that the frail results produced by survey-based analyses reflect in part the frail *measurement* of changing personal economic well-being. While relatively trustworthy measures of aggregate economic conditions are readily available, the reliability and validity of individual-level measures of personal financial security are largely untested. Most survey work, in fact, has depended on answers to a single question asked (with slight variation) by the National Elections Studies since 1952:

> We are interested in how people are getting along financially these days. Would you say that you (and your family living with you) are better off or worse off than you were a year ago?

Pinpointing the connection between economic circumstance and political preference based on this lone item is a perilous enterprise.[4]

Motivated by the theoretical issues at stake, we report here our efforts to improve the measurement of change in personal economic well-being. We attend to two kinds of measurement error: those arising from use of a single item rather than a set, and those deriving from use of a time frame (one year) that may be beyond respondents' memories. We draw upon a national survey conducted in the summer of 1983 to assess the adequacy of the traditional better/worse off question and compare its performance to new items we designed. While our results have immediate implications for the proper measurement of personal economic well-being, they also hold lessons for the measurement of subjective states more generally and for our understanding of the interplay between economics and politics in the lives of ordinary citizens.

---

[3] Kramer (1983), in addition, has argued that individual-level findings are statistical artifacts and that the mechanisms linking economic conditions to politics can never be uncovered using cross-sectional data.

[4] The better/worse off financially question is sometimes supplemented by others: by measures of unemployment and underemployment (e.g., Fiorina, 1981; Kinder and Kiewiet, 1979), by questions that invite citizens to name their most pressing personal problems, which are often economic (e.g., Kinder and Kiewiet, 1981; Kiewiet, 1983), and in at least two instances by sets of questions that attempt to map personal economic experiences and judgments more comprehensively (Kinder and Mebane, 1983; Mebane, 1982).

Measuring Change in Personal Economic Well-Being

THE CONCEPT

Because we cannot observe change in personal economic well-being directly, we asked respondents to report on perceptions and activities that change as a consequence of fluctuations in their own well-being. We assume each perception or activity—which is observable—is an indicator of change in economic well-being. We assume further that a person may cope with economic adversity in any of five ways. She can consume less or buy less expensive goods, look for more work, borrow, save less or dip into savings, or put off paying creditors. As her economic well-being improves, on the other hand, she can consume more, work less, forgo borrowing, save more, and pay debts promptly. Our survey questions tap each of these consequences of change in well-being. We also ask for a global assessment of whether family income has increased or decreased. Finally, we expand the standard better/worse off financially question by adding a follow-up that asks people how much better or worse off they have become.[5]

TIME FRAME

The traditional better/worse off financially question asks people to compare their current situation to last year's. Two considerations motivate us to consider a shorter time frame in this item as well as in our new ones. First, responses to questions asking for six-month instead of one-year comparisons may be more reliable. People rapidly forget even tangible and salient events like hospitalizations, automobile accidents, crime victimizations, and major home improvement expenditures (Cannell et al. (nd.); Penick, 1976; Neter and Waksberg, 1964). About *half* such instances are not reported to survey interviewers a year after their occurrences. Second, short-term changes in personal economic well-being may explain political preferences and behavior better than changes over a longer term. In evaluating presidents or deciding how to vote, people may ask, "What have you done for me lately?" and really mean lately.[6]

---

[5] Our conception of personal economic well-being does not include specific reference to unemployment. Unemployment is conceptually distinct from change in personal economic well-being. The two are related, to be sure, but a family can experience unemployment without experiencing economic adversity and can experience economic adversity without experiencing unemployment.

[6] Fair (1978) found that a six-month change in real GNP per capita is a slightly better predictor of presidential voting than a one-year change; Rosenstone (1982) found that short-term changes in unemployment suppress turnout more than long-term changes; Kernell (1978) found that six-month moving averages for differences in unemployment and consumer prices best explain fluctuations in presidential popularity.

ALTERNATIVE MEASURES

The survey questions we devised were included in the Center for Political Studies Pilot Survey for the 1984 National Election Study. The 314 people interviewed in July 1983 and reinterviewed a month later were a subsample of respondents to the 1982 National Election Study, a random sample of citizens of voting age.[7] To test the relative merits of the two time frames, respondents were randomly assigned to one of two groups. Each group was asked the same personal well-being questions, with one difference. People in sample A ($N = 158$) were prompted with phrases like "in the last year" or "this past year" while those in sample B ($N = 156$) were prompted with "in the last six months" or "this past six months."

The questions and the marginals for the two samples are reported in Table 1.[8] The very few people (.1 to 2.0 percent) who responded "don't know" or whose answers were not ascertained are deleted from the percentages displayed in the table. In general, the six-month time frame elicited slightly "better off" responses than the one-year prompt. This is to be expected: compared to the previous year, the economy really was better in the six months prior to the interviews. The rate of growth in per capita real disposable income between January 1983 and July 1983 was five times greater than between July 1982 and July 1983 (245 vs. 55 percent). Unemployment dropped .9 percent in the six months prior to the interviews, but fell only .3 percent over the previous year. Thus, it seems respondents did indeed match their recall to our time frames. We discuss other evidence on the discriminant validity of the two frames later.

MEASUREMENT MODEL

We employ the familiar Jöreskog (1970) maximum likelihood factor analysis model available in LISREL V to uncover the dimensions—perhaps more than one—that underlie our various measures of change in economic well-being. (For convenience, we coded all variables to a zero-one interval, with 1.0 representing the "better-off" responses.) Because we need to compare parameter estimates for the one-year and six-month time frames, we analyzed the two samples separately, estimating the coefficients from the variance-covariance matrix among the indicators.

---

[7] These data, as well as a description of sampling methods, are available from the ICPSR.

[8] The variable numbers reflect the order in which the questions were asked. Items 1–4 appeared as the opening questions on the questionnaire; the rest were asked three minutes later. Between the two sets of items respondents were asked about the national economy.

Table 1. Survey Questions Asked to Measure Changes in Personal Economic Well-Being—Variable Number, Question, and Marginals (in Percents)

Variable Number	*Question*	*Sample A 1 Year Time Frame*	*Sample B 6 Month Time Frame*
V1	We are interested in how people are getting along financially these days. Would you say that you (and your family living with you) are better off or worse off than you were (a year ago/six months ago)?		
	Better off	30.4	32.1
	Same	32.9	39.7
	Worse off	36.7	28.2
V2	Is that much better off or somewhat better off? Is that much worse off or somewhat worse off?		
	Much better off	7.6	7.1
	Somewhat better off	22.8	25.0
	Same	32.9	39.7
	Somewhat worse off	26.6	21.2
	Worse off	10.1	7.1
V3	During the (past year/past six months) have you (and your family) been able to buy most of the things you needed and planned on, or have you had to put off buying these things?		
	Had to put off buying	53.2	51.6
V4	Do you think that over the (last year/past six months) (your/your family's) income has gone up more than the cost of living, has it fallen behind, or has it stayed about even with the cost of living?		
	Gone up more	10.1	8.5
	Stayed about even	45.6	47.4
	Fallen behind	44.3	43.8
	Now I'm going to read a list of things that may have happened to you (or family members living with you) during the past (year/six months).		
V5	(This past year/In the last six months) have you (or anyone in your family) had to watch your budget more closely than the (year before/six months before)?		
	Yes	63.9	55.1
V6	During the past (year/six months), have you (and your family) put off medical or dental treatment because you didn't have the money?		
	Yes	29.7	27.6
V7	(This past year/In the past six months), in order to make ends meet, did (any of) you borrow money from a bank, lending institution, or from relatives or friends?		
	Yes	25.9	18.1

Table 1. *(Continued)*

Variable Number	*Question*	*Sample A 1 Year Time Frame*	*Sample B 6 Month Time Frame*
V8	[If yes] Did you have to borrow more, the same amount, or less than the (year/six months) before or didn't you have to borrow then?		
	More	3.8	6.5
	Same	5.7	3.9
	Less	5.7	1.3
	Didn't borrow then	10.8	6.4
V9	[Change in borrowing] [Constructed from V7 and V8]		
	Not borrowing	74.1	81.9
	Borrowing less	5.7	1.3
	Borrowing same	5.7	3.9
	Borrowing more	14.6	12.9
V10	(This past year/In the past six months) have you (and your family) had to use your savings to make ends meet?		
	Yes	47.5	37.8
V11	[If yes] Did you use your savings more, the same, or less than the (year before/six months before), or didn't you have to use your savings then?		
	More	30.4	12.3
	Same	5.7	13.5
	Less	1.9	2.6
	Didn't use savings then	9.5	9.0
V12	[If no] Over the past (year/six months) have you (and your family) been able to put money aside?		
	Yes	31.2	37.2
V13	[If yes] Have you been able to save more, the same amount or less than (the year before/during the six months before)?		
	More	10.2	10.3
	Same	15.9	20.5
	Less	5.1	6.4
V14	[Change in savings] [Constructed from V10, V11, V12, and V13]		
	Less savings	45.2	27.7
	Same savings	42.7	59.4
	Greater savings	12.1	12.9
V15	(This past year/In the past six months), in order to make ends meet, have (any of) you looked for a new job, or looked for a second job, or tried to work more hours at your present job?		
	Yes	44.9	34.2

As the estimates reported in Table 2 show, two distinct but correlated dimensions underlie the pool of indicators.[9] The first dimension produces economic evaluations: answers to the better/worse off financially, change in income, watch budget more closely, and change in savings questions. The remaining items—putting off buying things, putting off medical treatment, changing borrowing habits, and looking for a new job or more hours—are all reports of activities performed by people adjusting to economic changes. The evaluative dimension is comprised of things that people *think* about as a result of their financial states; the behavioral dimension is comprised of things that people *do* as a consequence of their financial states. (For similar results see Kinder and Mebane, 1983.) This model fits the data well regardless of whether the six-month or one-year time frame is employed. The adjusted measure of goodness of fit is .937 in both samples; the Chi-square has a probability of .657 in sample A and .638 in sample B.[10]

There is surely nothing in the measurement model that allows us to choose between the six-month and one-year time frames. The structure of the measurement model is precisely the same across the two samples; the loadings are nearly the same; the individual items are of comparable reliability; the two latent variables are correlated to the same degree; the estimated reliabilities of the indicators are virtually identical. We must search elsewhere for a basis to choose between the two time frames. But looked at in a more positive way, the results displayed in Table 2 provide striking confirmation of the robustness of the measurement model: the two dimensions of change in personal economic well-being show up in exactly the same manner in each of the two independent samples.

Construct Validity

We now have in hand four ways to measure personal economic well-being: the traditional three-category better/worse off financially ques-

---

[9] A blank cell indicates that the coefficient is statistically equivalent to zero. We initially treated each survey question as an indicator of a single latent dimension, but this specification fits the data poorly: prob. = .042 in sample A; prob. = .031 in sample B.

[10] We tested some alternative specifications. One allowed the savings question (V14) to load on both the evaluative and behavioral dimensions, but the item is apparently more an evaluation than a behavior. Saving does not really require that a person *do* anything; it is simply what is left in the checkbook at the end of the month. Borrowing, by contrast, involves a conference with a bank loan officer or a rich uncle. There was also no evidence of a response set within each of the two batteries of questions (V2–V4 and V5–V15). Finally, because the eight indicators are discrete (each has two to five response categories), we estimated the parameters using polychoric correlations, but found that the structure of the model and the relative loadings of the variables remained about the same. For details, see Rosenstone et al., (1983).

Table 2. Maximum Likelihood Factor Analysis of Items Measuring Changes in Personal Economic Well-Being (Estimates Made from Variance-Covariance Matrix)

		Loadings				
		Evaluative Dimension		*Behavioral Dimension*		
Variable	*Question*	*Coef*	*SE*	*Coef*	*SE*	*Reliability*
Sample A—1 Year Time Frame						
V2	Better/worse off financially	.193	(.021)			.495
V3	Put off buying things			.334	(.041)	.446
V4	Change in income	.251	(.024)			.588
V5	Watch budget more closely	.337	(.037)			.490
V6	Put off medical treatment			.310	(.038)	.457
V9	Change in borrowing			.163	(.032)	.193
V14	Change in savings	.214	(.027)			.393
V15	Look for new job/more hours			.264	(.042)	.281
						Total .898
Correlation between the two dimensions = .781						
Chi-square with 19 d freedom = 16.01 (prob = .657)						
Adjusted goodness of fit = .937						
Sample B—6 Month Time Frame						
V2	Better/worse off financially	.184	(.020)			.527
V3	Put off buying things			.378	(.042)	.567
V4	Change in income	.214	(.025)			.457
V5	Watch budget more closely	.366	(.038)			.538
V6	Put off medical treatment			.273	(.038)	.370
V9	Change in borrowing			.164	(.031)	.217
V14	Change in savings	.172	(.026)			.305
V15	Look for new job/more hours			.193	(.042)	.165
						Total .894
Correlation between the two dimensions = .759						
Chi-square with 19 d freedom = 16.29 (prob = .638)						
Adjusted goodness of fit = .937						

tion, the five-category version of the same question, the evaluative dimension, and the behavioral dimension.[11] We analyze two aspects of the convergent validity of each. First, how well is each predicted by

---

[11] We constructed a scale for the evaluative dimension and one for the behavioral dimension. The individual items were weighted by the factor scores generated from analyzing the variance-covariance matrix among the items comprising that dimension. (This was done separately for each dimension within each sample.) The scales are highly reliable:

	Sample A (1-Year)	Sample B (6-Month)
Reliability of evaluative scale	.81	.78
Reliability of behavioral scale	.70	.69

variables, like objective economic circumstances (changes in income, job loss) and social location (education, race, sex), that *should* be associated with changes in economic security? Second, do these associations vary in sensible ways with the time frames? To find out, we regressed each of the four measures of personal economic well-being on these validating variables and examined the unstandardized bivariate regression coefficients.

We found strong convergent validation for all four measures. Consider just a few examples. (For details see Rosenstone et al., 1983.) First, regardless of the measure of change in financial well-being used or the time frame employed, men are better off than women, whites better off than blacks, and the well educated are better off than those with less formal schooling. Second, unemployment in July 1983 is more highly associated with change in personal economic well-being measured over the last six months than over the last year. Third, change in family income between November 1982 and July 1983 is more strongly associated with one-year measures of change in personal economic well-being than with six-month measures.

One striking feature of these results is how well the traditional better/worse off financially item stands up to its competitors. The relationships between the various measures of objective economic circumstance and social location, on the one hand, and personal economic well-being, on the other, are about as strong when the latter is measured by the usual three-category format as when it is measured either by the five-category rendition or by the four items comprising the evaluative dimension scale. The traditional question is a respectably valid measure of change in personal economic well-being.

This point is reinforced by one final bit of validating evidence we gleaned from monthly data gathered by the Survey of Consumer Attitudes. We regressed responses to the traditional three-category better/worse off financially question with a one-year time frame onto monthly changes in aggregate per capita real disposable income. We estimated a series of equations, varying the number of months over which we calculated the change in real income. If people really do think back a full year as prompted, responses should be most highly associated with 12-month changes in income. As it happens, replies to the better/worse off financially questions are best predicted by changes in per capita real disposable income that are slightly longer than one year. The fit is best, and the coefficient is largest at about 14 months. (For details see Rosenstone et al., 1983.) Thus, people do seem to employ a baseline that is pretty close to the one they are asked to use.

In summary, the various measures of change in personal economic well-being, including the traditional three-category better/worse off financially question, perform well in our first set of validation tests. Each is predicted in expected ways by measures of objective economic circumstance and location in the social order. Moreover, various results suggest that respondents who are prompted to think back over the previous six months in fact do so; those prompted with a one-year frame look back over something close to one year. Personal economic well-being can indeed be measured adequately—a most important point—but these results do not single out the best measure among the several competitors. For that we must investigate the political consequences of personal economic well-being, our next and final test.

Predictive Validity

Here we examine, across the two time frames, how well the alternative measures predict three dependent variables: evaluations of the nation's economy, evaluations of Ronald Reagan's performance as president, and vote choices in a Reagan–Mondale trial heat (Tables 3–5).[12] For each dependent variable, we estimate a series of equations. The first equation employs the three-category better/worse off financially question as the measure of change in economic well-being; the second equation uses the five-category version; the third, the evaluative dimension; and the fourth, both the evaluative and the behavioral dimensions. The control variables are listed in a note at the bottom of each table.[13] We treat the financial well-being variables as endogenous and use two-stage least squares (2SLS) to correct for measurement error.[14]

We estimate a separate coefficient for the effect of change in personal economic well-being in each sample. By comparing the one-year

---

[12] We also estimated the effect of economic well-being on thermometer evaluations of Senator John Glenn to test the discriminant validity of the four measures. We expected the coefficients to be zero. They were.

[13] These demographic and political variables can safely be regarded as exogenous. We employed November 1982 measures of party identification, liberalism-conservatism, union membership, and family income to eliminate concerns about mediating effects of retrospective evaluations (Fiorina, 1981). The control variables are constrained to have the same effect in both samples.

[14] The causes of each measure of personal economic well-being discussed in the last section and the responses to the better/worse off financially question given in the November 1982 survey are used as exogenous instruments. The reduced form equations have sufficiently high R^2s—.29 to .44 for endogenous variables constrained to be equal across equations, .59 to .82 for variables unconstrained across equations—to yield reliable second-stage estimates.

Table 3. Estimated Effect of Alternative Measures of Changes in Personal Economic Well-Being on Evaluations of the Nation's Economy—Two-Stage Least Squares Estimates[a]

Eq #	*Measure of Change in Personal Economic Well-Being Appearing in the Equation*	*Coefficients* Sample A (1 Year)	*Coefficients* Sample B (6 Months)	*Prob. Coef A = Coef B*	*Standard Error of Regression*
1.	Better/worse off financially—3 categories	.135 (.076)	.214 (.074)	.03	.224
2.	Better/worse off financially—5 categories	.228 (.110)	.314 (.106)	.06	.221
3.	Evaluative dimension	.294 (.113)	.337 (.105)	.25	.221
4.	Evaluative dimension	.080 (.200)	.355 (.279)		.224
	Behavioral dimension	.240 (.145)	.099 (.184)		
5.[b]	Evaluative dimension	.280 (.180)			.222
	Behavioral dimension	.136 (.118)			
6.[c]	Evaluative dimension	.433 (.218)	.549 (.213)		.236 (A) .212 (B)
	Behavioral dimension	−.038 (.191)	−.066 (.151)		

[a] The other variables that appeared in each equation were: party identification in November; liberalism-conservatism in November; log (family income in November); and Hispanic. These variables are assumed to be exogenous. (In addition, dummy variables for rural residents, women, and unemployed respondents or mates appear as exogenous variables in equations 1 and 2, but were deleted [prob. < .65] from equations 3–6. Whether the respondent or his mate was unemployed during the year appeared as a variable in equations 1–4, but was deleted from equations 5 and 6 for the same reason.) The measures of personal economic well-being are treated as endogenous.

[b] In this equation, the coefficients were constrained to be equal across the two samples.

[c] These coefficients were estimated separately on each sample.

and the six-month coefficients, we can assess the predictive power of each question format. To help us in this task, we report the probability that the two coefficients are equal.[15]

---

[15] The coefficients in the fourth equation for each dependent variable—that which pits the behavioral dimension against the evaluative dimension—are too unstable to allow a meaningful comparison of the relative effects of the two dimensions. We are forced, therefore, to assume that the effects are constant across the two samples; these estimates are reported in equation 5. Finally, because evaluations of the nation's economy also were asked with both a one-year and a six-month time frame, it may be inappropriate to assume that the effects of the other exogenous variables are equal in the two samples. For this variable, then, we estimate separate equations for each sample (reported in equation 6, Table 3 only).

Table 4. Estimated Effect of Alternative Measures of Changes in Personal Economic Well-Being on Evaluations of Reagan's Handling of His Job as President—Two-Stage Least Squares Estimates[a]

		Coefficients			
Eq #	*Measure of Change in Personal Economic Well-Being Appearing in the Equation*	*Sample A (1 Year)*	*Sample B (6 Months)*	*Prob. Coef A = Coef B*	*Standard Error of Regression*
1.	Better/worse off financially—3 categories	.436 (.114)	.323 (.104)	.22	.303
2.	Better/worse off financially—5 categories	.584 (.155)	.479 (.144)	.09	.303
3.	Evaluative dimension	.698 (.166)	.525 (.139)	.03	.300
4.	Evaluative dimension	.489 (.310)	.639 (.298)		.301
	Behavioral dimension	.116 (.206)	−.079 (.207)		
5.[b]	Evaluative dimension	.607 (.238)			.303
	Behavioral dimension	.007 (.161)			

[a] The other variables that appeared in each equation were: party identification in November; liberalism-conservatism in November; union household in November; log (family income in November); education in November; race; Jew; and age. These variables are assumed to be exogenous. (Whether the respondent or his mate was unemployed in the last year also appeared as an exogenous variable in equations 1 and 2, but was deleted from equations 3–5 because its probability of being different from zero fell to .65.) The measures of personal economic well-being are treated as endogenous.

[b] In this equation, the coefficients were constrained to be equal across the two samples.

THREE VS. FIVE CATEGORY BETTER/WORSE OFF FINANCIALLY QUESTION

Our results are unequivocal and entirely startling. Simply by expanding the traditional measure's three response categories to five, we find sizable increases in estimated political effects. The increases show up in *all* comparisons: for evaluations of the national economy, ratings of President Reagan's performance, and vote choices in a Reagan–Mondale trial beat, and for six-month and one-year time frames alike. Put the other way around, in these data, *employing the traditional better/worse off financially question causes the political effect of changing personal economic conditions to be underestimated by as much as one-half.* For example, the estimated impact of changes in personal economic well-being on evaluations of Reagan's performance as president, after correcting for measurement error, is .436 using the

Table 5. Estimated Effect of Alternative Measures of Changes in Personal Economic Well-Being on Reagan-Mondale Trial Heat—Two-Stage Least Squares Estimates[a]

Eq #	*Measure of Change in Personal Economic Well-Being Appearing in the Equation*	*Coefficients*		*Prob. Coef A = Coef B*	*Standard Error of Regression*
		Sample A (1 Year)	*Sample B (6 Months)*		
1.	Better/worse off financially—3 categories	.462	.424	.27	.397
		(.157)	(.152)		
2.	Better/worse off financially—5 categories	.627	.612	.44	.392
		(.220)	(.215)		
3.	Evaluative dimension	.463	.421	.37	.386
		(.230)	(.205)		
4.	Evaluative dimension	.824	.317		.394
		(.446)	(.430)		
	Behavioral dimension	−.226	.059		
		(.297)	(.287)		
5.[b]	Evaluative dimension	.503			.389
		(.330)			
	Behavioral dimension	−.055			
		(.214)			

[a] The other variables that appeared in each equation were: party identification in November; liberalism-conservatism in November; union household in November; log (family income in November); education in November; and Hispanic. These variables are assumed to be exogenous. The measures of personal economic well-being are treated as endogenous.

[b] In this equation, the coefficients were constrained to be equal across the two samples.

three-category measure but .584 when the five-category version is adopted (Table 4).[16]

EVALUATIVE DIMENSION VS. THE BETTER/WORSE OFF FINANCIALLY QUESTIONS

The evaluative dimension sometimes outperforms the five-category version of the better/worse off financially question and sometimes does not. It always outperforms the traditional three-category format. Thus, the marginal gain from employing the four-item scale over the five-category better/worse off financially item is not nearly as great as the gain from using the five- instead of the three-category better/worse off financially question.

---

[16] Had we failed to correct for measurement error, the coefficients, particularly for the traditional measure, would of course be attenuated still further. If, for example, we had used ordinary least squares instead of 2SLS, the coefficients for the one-year time frame in the first three equations listed in Table 4 would have been .271, .365, and .455 instead of .436, .584, .698.

THE EVALUATIVE VS. THE BEHAVIORAL DIMENSION

The behavioral dimension has no independent explanatory power over and above that of the evaluative dimension (look at equations 4 and 5 in each table).[17] This finding is not an artifact of the high association between these two dimensions; relatively efficient estimates are produced once the effects in the two time frames are constrained to be equal (in equation 5).[18]

THE SIX-MONTH VS. THE ONE-YEAR TIME FRAME

The contest between the two time frames does not produce a decisive winner. The six-month time frame is more strongly associated with evaluations of the nation's economy; the one-year time frame is more strongly related to evaluations of Reagan's performance as president, and choices between Reagan and Mondale. Few of these differences are sure bets: most of the probabilities fall short of conventional standards of statistical significance.

Conclusions

Where others have proposed theoretical solutions to the puzzling divergence between aggregate and individual-level findings on the political consequences of economic fluctuations, we have proposed a methodological one. The single survey item traditionally used to prove (or disprove) the link between changing economic well-being and political evaluations and behaviors, although valid, is quite unreliable relative to an expanded version of the same question and relative to the evaluative scale built from four items. The empirical consequences in our data are striking: with new measures, changes in respondents' personal economic well-being exercise substantial influence over their assessment of the national economic scene, the support they give their president, and vote choice.[19] Reliance on the traditional item alone seriously underestimates the political impact of changing personal economic well-being. In some degree, then, our analysis illustrates the perilous consequences that follow from the violation of an altogether familiar maxim. *Items that produce greater variance in responses are preferred to items that produce less variance.* Surely, our most startling result is the sharp boost in predictive validity that occurred when the five-category version of the better/worse off question replaced the traditional three-category version.

---

[17] This conclusion also holds when the five-category better/worse off financially question is used in lieu of the evaluative dimension in equation 5 in Tables 4 and 5, and in equation 6 in Table 3.

[18] We suspect that the behavioral dimension is a cause of the evaluative dimension. The behavioral scale is more strongly associated with unemployment than either the worse off financially questions or the evaluative scale. Moreover, variables measuring experience with unemployment drop out when the behavioral dimension is added to the equations for evaluations of the nation's economy and evaluations of Reagan's job performance.

[19] Sears and Lau (1983) argue that the political effect of change in personal economic well-being stems from asking respondents about their economic circumstances either immediately before or after soliciting their political judgments. Interviewers either inadvertently prime respondents with personal economic reasons for their political judgments, or respondents rationalize their political judgments in personal economic

Another striking result is the relatively poor performance of the behavioral dimension. Perhaps this is because the opportunities to engage in the activities that comprise the behavioral dimension vary across the population. Not all people cope with economic fluctuations in the same way. The elderly, we discovered, are less likely to put off buying things than to postpone medical and dental care. Blacks, perhaps reacting to realistic assessments of the job market, are less likely to look for work than to borrow, forgo medical treatment, or defer purchases. The variables that make up the evaluative dimension, on the other hand, are not subject to selective opportunities. Restriction of variability to certain subgroups may account for the behavioral dimension's lower reliability and weaker explanatory power.

It may also be that questions that ask for recall of specific instances are generally less reliable than questions that ask for comparisons of current states to past baselines. An anecdote will make the point. Nineteen respondents had said in their November 1982 interviews that either they or their spouses were unemployed at that time. Eight months later, in July 1983, only 8 of the 19 (42 percent) said that they or their spouses were unemployed *in the past year*. Assuming that there is little overreporting of current unemployment, this is a stunning demonstration of the rapidity with which memories of specific facts or behaviors lapse, a finding well in keeping with other studies of recall cited earlier. Although people find it difficult to recall isolated events (as in the items comprising the behavioral dimension), our analysis suggests that people can capably draw comparisons with recalled baselines (as in the questions that make up the evaluative dimension).

Although our results do not settle the theoretical debates that served as our point of departure, they do send some important warnings to the

---

terms. Our results stand up well to their arguments. First, 2SLS purges the coefficients of any effect of rationalization. Second, although questions about the national economy and Reagan's performance were asked in close proximity to the personal economic well-being questions, the Reagan–Mondale trial heat question appeared in an altogether different interview: it led off the second wave, some four weeks later. Change in personal economic well-being is no worse a predictor of trial heat choice than of the other two variables.

protagonists. Surely the debates cannot be decided conclusively in the absence of better data. The construct "change in personal economic well-being" is but one character in the theoretical cast; better measures of "change in national economic well-being" and "change in group economic well-being," to name two, must be developed as well. Nor is adequate measurement the whole battle. One customary and likely consequential difference between aggregate and survey studies of political economy is that the economic context varies in the first but not the second (Kramer, 1983). When over-time variance is introduced into survey-based studies along with adequate measures of personal, group, and national economic well-being, then the individual-level mechanisms that link economic changes to political outcomes may be more convincingly demonstrated.

References

Bloom, H. S., and H. D. Price
1975 "Voter response to short-run economic conditions: the asymmetric effect of prosperity and recession." American Political Science Review 69:1240–54.

Cannell, C. F., K. H. Marquis, and A. Laurent
n.d. "A summary of studies of interviewing methodology." Unpublished manuscript, Survey Research Center, University of Michigan.

Conover, P. J.
1985 "The impact of group economic interests on political evaluations." American Politics Quarterly 13:139–66.

Fair, R. C.
1978 "The effect of economic events on votes for president." Review of Economics and Statistics 60:159–73.

Feldman, S.
1982 "Economic self-interest and political behavior." American Journal of Political Science 26:446–66.

Fiorina, M. P.
1978 "Economic retrospective voting in American national elections: a microanalysis." American Journal of Political Science 22:426–43.
1981 Retrospective Voting in American National Elections. New Haven: Yale University Press.

Frey, B. S., and F. Schneider
1975 "On the modeling of politico-economic interdependence." European Journal of Political Research 3:339–60.

Hibbs, D. A., Jr.
1982 "The dynamics of political support for American presidents among occupational and partisan groups." American Journal of Political Science 26: 313–32.

Hibbs, D. A., Jr., and H. Fassbinder
1981 Contemporary Political Economy. Amsterdam: North-Holland.

Jacobson, G. C., and S. Kernell
1981 Strategy and Choice in Congressional Elections. New Haven: Yale University Press.

Jöreskog, K. G.
1970 "A general method for analysis of covariance structures." Biometrika 57:239–51.

Kernell, S.
1978 "Explaining presidential popularity." American Political Science Review 72:506–22.
Kiewiet, D. R.
1983 Macroeconomics and Micropolitics: The Electoral Effects of Economic Issues. Chicago: University of Chicago Press.
Kinder, D. R.
1981 "Presidents, prosperity, and public opinion." Public Opinion Quarterly 45:1–21.
Kinder, D. R., and D. R. Kiewiet
1979 "Economic discontent and political behavior: the role of personal grievances and collective economic judgments in congressional voting." American Journal of Political Science 23:495–527.
1981 "Sociotropic politics: the American case." British Journal of Politics 11:129–61.
Kinder, D. R., and W. R. Mebane, Jr.
1983 "Politics and economics in everyday life." In K. Monroe (ed.), The Political Process and Economic Change. New York: Agathon.
Kinder, D. R., S. J. Rosenstone, and J. M. Hansen
1983 "Group economic well-being and political choice." Pilot study report to the 1984 NES planning committee and NES board.
Kramer, G. H.
1971 "Short-term fluctuations in U.S. voting behavior, 1896–1964." American Political Science Review 65:131–43.
1983 "The ecological fallacy revisited: aggregate- versus individual-level findings on economics and elections and sociotropic voting." American Political Science Review 77:92–111.
Mebane, W. R., Jr.
1982 "The warp of sociotropic thinking." Paper delivered at the annual meeting of the Midwest Political Science Association, May 1982, Milwaukee, Wisconsin.
Neter, J., and J. Waksberg
1964 "A study of response errors in expenditures data from household interviews." Journal of the American Statistical Association 59:18–55.
Penick, B. K. E.
1976 Surveying Crime. Washington, DC: National Academy of Sciences.
Rosenstone, S. J.
1982 "Economic adversity and voter turnout." American Journal of Political Science 26:25–46.
1983 Forecasting Presidential Elections. New Haven: Yale University Press.
Rosenstone, S. J., J. M. Hansen, and D. R. Kinder
1983 "Measuring personal economic well-being." Report submitted to the board of overseers, National Election Study, October 1983.
Schlozman, K. L., and S. Verba
1979 Injury to Insult. Cambridge: Harvard University Press.
Sears, D. O., and R. R. Lau
1983 "Inducing apparently self-interested political preferences." American Journal of Political Science 27:223–52.
Sears, D. O., R. R. Lau, T. R. Tyler, and M. A. Harris, Jr.
1980 "Self-interest versus symbolic politics in policy attitudes and presidential voting." American Political Science Review 74:670–84.
Tufte, E. R.
1978 Political Control of the Economy. Princeton: Princeton University Press.

Race Differences in Abortion Attitudes

ELAINE J. HALL AND MYRA MARX FERREE

SINCE 1965 a consistent finding in public opinion surveys of abortion attitudes has been that black respondents are less in favor of legal abortion than white respondents. This is true whether one looks at the NORC General Social Surveys (Arney and Trescher, 1976; Granberg and Granberg, 1980; Evers and McGee, 1977) or at other surveys using differently worded questions (Westoff and others, 1969, using NFS; Lipson and Wolman, 1972, using a Population Commission survey; Pomeroy and Landman, 1972, using Gallup Polls). Only Combs and Welch (1982) take the further step and focus upon the variables which may explain this racial difference in prochoice attitudes.

Using the NORC surveys for 1972 to 1980, Combs and Welch attempt to explain racial differences in legal abortion support by three antecedents: (1) lower SES, as measured by family income and years of education, (2) stronger religious practice, as measured by frequency of church attendance, denominational membership (Catholicism and

---

Abstract Public opinion surveys since 1965 find that black respondents are less in favor of legal abortion than white respondents. Using the 1982 NORC General Social Survey, we replicate and expand one of the few studies (Combs and Welch, 1982) that examined the structure and determinants of prochoice attitudes of blacks and whites. Our major findings are (1) the racial difference in prochoice attitude is as great in 1982 as in the 1970s, (2) contrary to the suggestion of Combs and Welch, the demographic and attitudinal determinants of abortion attitudes differ for blacks and for whites, and (3) for those respondents who differentiate their acceptance of legal abortion, the pattern of prochoice attitudes also differs by race.

Elaine J. Hall is a doctoral student in the Department of Sociology, University of Connecticut, Storrs, and Myra Marx Ferree is an Associate Professor of Sociology in that department. The authors wish to thank Mark Abrahamson, Nancy Andes, and the anonymous reviewers for this journal for their constructive criticism of an earlier draft. The article was originally presented as a paper at the Eastern Sociological Society Annual Meetings, March 1985.

Public Opinion Quarterly Vol. 50:193–207 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-193/$2.50

fundamentalism), and the respondent's self-identified strength of affiliation, and (3) southern culture, measured by present region of residence and at age 16, and by present size of community and at age 16. Each of these variables has been shown to have a negative zero-order correlation with abortion support. Each can also be assumed to be more prevalent among blacks than whites. Therefore, Combs and Welch predict that controlling for these three antecedents would reduce or eliminate the racial difference in abortion attitudes.

In fact, Combs and Welch find that controlling for demographic variables (SES, southern culture, age and sex) reduces the difference between blacks and whites by only about one-third. When religious practice variables are also controlled, the black-white difference declines further. However, even when controlling for all 11 indicators, a small but statistically significant racial difference in abortion attitudes remains.

Although there is a statistically significant racial difference in each subperiod which Combs and Welch examine, the size of this difference appears to diminish over the 1970 decade. Combs and Welch therefore suggest that, with a few minor exceptions, "patterns in attitudes towards abortion in the black community mirror those in the larger society, and thus seem to be shaped largely by the same forces as white attitudes" (1982:517). In short, they see the differences between blacks and whites as essentially residual effects of historically different social positions that are converging over time.

Operating from a demographic perspective on abortion attitudes, Combs and Welch assume a longitudinal trend toward increased support of legal abortion in all groups as they become more urban, educated, and secular. These demographic trends are generally seen as producing more "modern" or progressive attitudes on a variety of attitudinal measures, of which acceptance of legal abortion is only one. In contrast, a feminist perspective assumes a conflict model of increasing polarization of attitudes between groups as the debate over women's equality and the family is politicized (Deitch, 1982; Mueller, 1983; Conover and Gray, 1983). This feminist explanation stresses attitudinal variables, measures of an underlying belief system regarding family and women's roles, as antecedents of abortion attitudes. From this perspective, availability of legal abortion is a central issue in obtaining reproductive freedom, and reproductive self-determination is a central issue in obtaining social, political, and economic equality for women (Luker, 1984).

Attitudinal variables suggested by the feminist perspective have been shown to be important predictors of abortion attitudes, even when demographic antecedents are controlled (Blake and DelPinal, 1980). Granberg and Granberg find that while demographic variables explain 10 percent of the variation in prochoice support in the 1972–80 NORC data, a personal morality index explains 21 percent of the variation. Attitudes regarding premarital sexual relations and ideal family size are the two strongest and more consistent predictors of abortion attitudes in this index (Granberg and Granberg, 1980:255–56). Deitch (1982) uses opinions on premarital sex, homosexuality, women's roles, liberal-conservative self-identification, and government spending priorities in an analysis of the 1972–80 GSS. She finds the associations between prochoice position and each of these attitudes do not disappear when education, sex, and religious affiliation are controlled. Weak positive correlations with prochoice are reported for both support for ERA (Granberg and Granberg, 1980:259) and for liberal-conservative attitudes toward economic and political roles for women (Granberg, 1978:422).

While the connection between position on legal abortion and other attitudes is important, it has never been explored separately for blacks. Other research on black attitudes toward feminism suggests that the determinants may be different for the two racial groups (Ransford and Miller, 1983). They find middle-class blacks more conservative in their attitudes towards women's roles than other blacks. This relationship is the reverse of that typically found for whites, and one that is not replicated on other non-sex-role issues, Their work also implies a different effect of gender in the two groups. Black men are found to be more traditional than white men, while black and white women are equally feminist when the demographic differences between them are controlled. Rather than the convergence anticipated by Combs and Welch's demographic model, the feminist perspective anticipates continued and even increased differences between blacks and whites if the demographic and attitudinal roots of a prochoice position are different in the two groups, as Ransford and Miller find the correlates of feminist attitudes to be. Moreover, Combs and Welch note that black respondents' attitudes toward abortion are less scalable than those of whites. This differential scalability may reflect differences of attitude structure in the two population groups that deserve further scrutiny.

The purposes of this article are to extend the analysis of Combs and Welch forward to the 1984 General Social Survey, to specify additional attitudinal determinants of the relative lack of support for legal abortion among blacks, and to compare the structure of white and black abortion attitudes. Specifically, we hypothesize the following:

1. The racial difference in abortion attitudes is persistent over time. We expect no evidence of a trend toward convergence when black and white attitudes are examined for the entire 1972–1984 GSS series.

2. The racial difference in prochoice attitudes is not due solely to positional antecedents. We expect a significant racial difference in prochoice attitudes to remain even when controlling for the demographic and religious practice variables examined by Combs and Welch.

3. The inclusion of attitudinal variables will enhance the explanatory power of the model but will not completely eliminate the difference by race. We expect both feminist and sexual morality attitudes to significantly increase the variance explained but to leave a statistically significant racial gap, expressing the political dimension of difference.

4. Because of the political component in prochoice position, we expect black and white attitudes to rest in part on different antecedents. This would be evident if an antecedent (such as gender) is significant for one race but not for the other, or if an antecedent (such as age) is significant for both races but varies in strength or direction of influence. Among these interactions, we expect feminist attitudes to be more predictive of a prochoice position for whites than for blacks, who may have conflicting political priorities.

5. The structure of abortion attitudes is itself different for blacks and whites. When support for a prochoice position is qualified by situation, we expect to find different patterns of partial support for blacks and whites.

Methods

Because of the small number of black respondents in each yearly NORC survey from 1972 to 1980, Combs and Welch combined their data into periods of three years each: 1972–74 (671 blacks), 1975–77 (468 blacks), and 1978–80 (298 blacks). Pooling data hides variation within each period; this plus the small number of black respondents in the last of these combined GSS samples makes it difficult to determine trends.[1] We expand the series to 1984 and break out each year of the GSS for trend analysis, but then concentrate on the 1982 GSS, where the oversampling of blacks eliminates the need to pool data to get an adequate number of cases for analysis of black attitudes. After excluding 27 respondents of "other" races, the total 1982 sample of 1833 consists of 28 percent (510) blacks and 72 percent (1323) whites. When black and white respondents are analyzed together, weighting is used to make the sample representative (Smith, 1984).

---

[1] The interpretation of trends across the three time periods is further complicated by sampling issues. The 1972–74 surveys were a modified probability design which used quotas for age, sex, and employment status on the block level. Because the 1972 GSS was drawn on the 1960 census information, the proportion of blacks in this survey is atypically high (261, or 16.2 percent of the 1613 total). In the 1975–77 period, problems of refusal and completion rates produced a low proportion of blacks in the 1976 GSS (only 129 or 8.6 percent of the 1499 total). Although a full probability sampling design was used in the 1978–80 period, the total number of respondents is limited to only two surveys (1978 and 1980). Overall, these sampling issues combine to give the effect of a diminishing number of black respondents in the three periods defined by Combs and Welch.

Following Combs and Welch, we construct a six-point additive scale of prochoice attitude by summing positive responses to each of the six questions about situations where abortion should or should not be legal. The demographic predictors are measured as follows: (1) SES by years of education and family income,[2] (2) gender and age of the respondent, (3) Southern culture by region presently and at age 16 (1 = South, 2 = non-South), and by size of community presently (1 = city, 2 = suburb, 3 = other) and at age 16 (1 = nonrural, 2 = rural), and (4) religious practice variables measured by frequency of church attendance, strength of affiliation, and dummy variables for Catholicism and fundamentalism. For the following attitudinal variables, the smaller the value, the more traditional the definition of women's roles: (1) extent of support for ERA, (2) attitude toward premarital sexual relations, (3) belief that women should stay home and leave running the country to men, and (4) belief that married women should not work. Attitude toward family size reflects the centrality of motherhood in the definition of women's roles; the larger the ideal family size, the more traditional the definition of women's roles.

Results

THE TREND ANALYSIS

First, we examine the racial difference in prochoice attitudes from 1972 to 1984 for both the additive prochoice scale and for each of the six reasons for an abortion. In Table 1 we note a substantial and statistically significant difference in prochoice attitudes between blacks and whites for each year. The mean scale scores reveal that whites are more prochoice for each of the 11 years, a difference that averages approximately three-quarters of a point on the six-point scale. Whites are also consistently more prochoice than blacks for each of the six reasons, although the difference is not significant for each reason for each year. The mean percentage difference per reason ranges from

---

[2] Income has been recoded to the category midpoint in dollars for each year. The income variable used in each survey was the one designed to reflect the income distribution at that point; with inflation, categories have changed. For details, see pages 48–51 in Davis and Smith (1982). In all other cases, the coding follows Combs and Welch (1982:512–13).

Table 1. Distribution of Abortion Support by Race, 1972–1984

	Health			*Rape*			*Defect*			*Can't Afford*			*Single*			*Wants No More*			*Prochoice Scale*		
Year	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*White*	*Black*	*Diff*
1972	86	70	16*	78	56	22*	78	56	22*	46	36	10*	43	32	11*	39	30	9*	3.71	2.76	.95*
1973	92	80	12*	83	66	17*	85	67	18*	54	38	16*	50	30	20*	48	36	12*	4.10	3.15	.95*
1974	91	87	4	84	71	13*	84	71	13*	54	42	12*	50	33	17*	46	36	10*	4.09	3.40	.69*
1975	89	82	7*	82	68	14*	82	63	19*	51	45	6	47	37	10*	44	40	4*	3.95	3.38	.57*
1976	90	81	9*	82	67	15*	83	65	18*	52	41	11*	50	33	17*	45	38	7	4.01	3.22	.89*
1977	90	77	13*	83	65	18*	86	67	19*	53	43	10*	49	35	14*	45	42	3	4.04	3.22	.82*
1978	89	85	4*	83	65	18*	82	67	15*	46	40	6	40	33	7	39	36	4	3.78	3.25	.53*
1980	89	81	8*	82	70	12*	82	66	16*	50	53	7	48	36	12*	46	40	6	3.96	3.34	.62*
1982	90	82	8*	85	68	17*	83	66	17*	50	39	11*	48	31	17*	47	34	13*	4.02	3.19	.83*
1983	88	82	6*	80	72	8*	78	66	12*	42	39	3	39	26	13*	38	35	3*	3.58	3.15	.43*
1984	89	81	8*	80	63	17*	80	63	17*	46	35	11*	45	30	15*	43	30	13*	3.80	3.01	.79*
Mean percentage diff			8.64*			15.55*			16.91*			9.36*			13.91*			7.55*			
Mean item score	1.11	1.20		1.18	1.34		1.18	1.35		1.50	1.60		1.54	1.68		1.56	1.64		3.91	3.17	.74*
Race-year interaction (*F*)	2.58 ($p < .005$)			1.35 ($p > .20$)			.82 ($p > .61$)			.72 ($p > .71$)			.88 ($p > .55$)			1.03 ($p > .55$)			1.19 ($p > .29$)		

* Significant at .05 level.

almost 9 to 17 percentage points. The smallest race difference is found in the item where whites are least supportive: when the woman is married and wants no more children.

Second, the linear trend of diminishing differences reported by Combs and Welch is not supported in these data. Despite fluctuations in the attitudes of both whites and blacks throughout the period, the most meaningful change is an increase in support by both groups between 1972 and 1973, probably reflecting the Supreme Court's legalization of abortion at that time. There is no consistent trend toward convergence between the two groups for either the individual reasons or the prochoice scale for the full 1972–84 period or for the more restricted set of 1972–80 data that Combs and Welch used. For both of these periods, the *F* statistics for the year–race interaction are not significant for either the prochoice scale nor, with one exception, for the individual reasons.[3] Looking at abortion when the mother's health is in danger, the one item that shows significant convergence, we find that with the exception of 1972, the white marginals can be represented as 90 ± 2 and those of blacks as 82 ± 5. The interaction thus seems to show that in aftermath of *Roe v. Wade* on this item, blacks increased their support more than whites, who were already very supportive. For the other items, whites as well as blacks increased their approval, so the magnitude of the difference did not decline.

Looking at the prochoice scale in Table 1, we can see how pooling data into three periods (1972–74, 1975–77, and 1978–80), as Combs and Welch did, could create the appearance of a trend toward disappearing differences. The years of least difference (1978 and 1983), however, are notable for their atypically low levels of white approval of legal abortion for the more discretionary (often called "soft") reasons, rather than for increases in black support. Overall, we can conclude that the stability of abortion attitude is much more impressive than any changes for either blacks or whites.

EXPLAINING THE RACE GAP

Turning now to attempts to explain this persistent race difference, we see in Table 2 that the demographic and religious practice variables employed by Combs and Welch continue to have about the same effect on the impact of race on abortion attitudes as they did in earlier years. In 1982, the unstandardized regression coefficient (−.83) for race is substantially reduced (to −.71) with the control of the demographic variables, and is reduced even further (to −.55) but remains significant

---

[3] Combs and Welch also note that the trend toward convergence in their data is not significant using a chi-square goodness of fit test (Taylor, 1980).

Table 2. Impact of Race on Abortion Attitudes, 1972–1984[a]

	1972–1974			*1975–1977*			*1978–1980*			*1982*			*1983–1984*		
Variables	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2
Race alone	−.90*	(−.15)	.02	−.70*	(−.11)	.01	−.57*	(−.09)	.01	−.83*	(−.13)	.02	−.60*	(−.09)	.01
Race with controls for demographic variables[b]	−.49*	(−.08)	.12	−.43*	(−.07)	.09	−.32*	(−.05)	.08	−.71*	(−.11)	.13	−.42*	(−.07)	.08
Race with controls for demographic & religious practice variables[c]	−.45*	(−.08)	.24	−.27*	(−.04)	.23	−.20	(−.03)	.23	−.55*	(−.09)	.25	−.24	(−.04)	.22
N[d]	4054			4019			2659			1604			2627		

* Significant at .05 level.

[a] Since this article does not intend to be an exact replication of Combs and Welch, we code some variables differently than they did. We use XNORCSIZ instead of SRCBELT to measure present size of community; either variable is coded 1 = city, 2 = suburb, and 3 = other. To allow for inflation, four income variables are used to cover the 1972–1984 period (INCOME72, INCOME, INCOME77, and INCOME82). All income variables are coded to the midpoint of the category. Frequency of church attendance is coded to number of weeks per year to make it a legitimately interval variable.

[b] Demographic variables: education, income, age, sex, region, size of community, region and size of community at age 16.

[c] Religious practice variables: attendance, Catholicism, fundamentalism, strength of affiliation. Strength of affiliation is not available for the 1972–1974 period and therefore is not included in the regression for that period.

[d] These results are based on pairwise deletion. The regressions were also conducted using listwise deletion, revealing no substantial differences. The 1982 regression was weighted to correct for the oversample of blacks.

when religious practice variables are controlled. The size of the residual gap is thus roughly half a point on a six-point scale. The same pattern occurs for the 1983–84 period (a coefficient of −.60 is reduced to −.42 and further to −.24) except that the racial difference with all variables in the regression is no longer significant, contrary to our hypothesis. It is difficult to interpret this as a trend, however, since the 1983 difference is at least as large as the difference in 1972–74. The substantial amount of variation explained by demographic and religious practice variables during the 1970s (adjusted R^2 of .24, .23, .23 for the three pooled time periods) continues in the 1980s (adjusted R^2 of .25 and .22).

In the years for which these measures are available, the inclusion of attitudinal variables as antecedents to abortion attitudes further enhances the explanatory power of the model but does not diminish the racial differences in prochoice attitudes. Four attitudinal variables—premarital sex, ideal family size, women working, and women staying out of politics—were asked in the 1975, 1977, and 1982 General Social Surveys. Adding them as a fourth step to the regression used in Table 2 explains an additional amount of the variance in abortion attitudes. The adjusted R^2 increases from .21 to .28 in 1975, from .24 to .32 in 1977, and from .25 to .34 in 1982. In all three years when this set of attitude variables is available, controlling for feminism and fertility attitudes actually increases the residual difference by race. The size of the gap varies from −.25 in 1975 to −.46 in 1977 and −.66 in 1982, thus showing a slight increase over this period.

Further examination of the antecedents of black attitudes toward

Table 3. The Impact of Demographic and Attitudinal Variables on Abortion Attitudes for Blacks and for Whites, 1982

	Blacks					*Whites*					*F (race interaction)*
		Demographic		*Attitudinal*			*Demographic*		*Attitudinal*		
	Simple r	*b*	*(Beta)*	*b*	*(Beta)*	*Simple r*	*b*	*(Beta)*	*b*	*(Beta)*	
Income	.21*	.013	(.08)	.013	(.08)	.19*	.012*	(.08)	.010*	(.07)	0.96
Education	.32*	.085*	(.14)	.059	(.10)	.26*	.114*	(.18)	.091*	(.14)	0.20
Sex	.00	.364	(.09)	.425*	(.10)	−.03	.101	(.03)	.178	(.05)	1.44
Age	−.29* +	−.013*	(−.12)	−.012* +	(−.10)	−.11* +	.004	(.04)	.016* +	(.15)	10.08*
Size of residence	−.19*	−.169	(−.08)	−.144	(−.06)	−.15*	−.175*	(−.08)	−.128*	(−.05)	0.07
Region	.24* +	.101	(.02)	−.115	(−.00)	.10* +	−.147	(−.03)	−.228	(−.05)	0.89
Rural residence 16	−.29*	−.526*	(.12)	−.391	(.09)	−.20*	−.312*	(.07)	−.155	(.04)	0.42
Region 16	.30* +	.458	(.10)	.486*	(.11)	.11* +	.125	(.03)	.016	(.00)	1.34
Catholicism	.04	−.138	(−.02)	−.159	(−.02)	−.06*	−.384*	(−.09)	−.382*	(−.08)	0.17
Fundamentalism	−.10* +	−.002	(−.00)	.083	(.02)	−.22* +	−.381*	(−.08)	−.142	(−.03)	0.30
Attendance	−.26* +	−.018*	(−.22)	−.009*	(−.10)	−.38* +	−.025*	(−.34)	−.014*	(−.19)	0.51
Premarital sex	.37* +			.352* +	(.23)	.48* +			.529* +	(.34)	4.90*
Ideal family size	−.26*			−.121*	(−.12)	−.23*			−.177*	(−.13)	0.21
Pro-ERA	.10* +			.252*	(.11)	.23* +			.215*	(.10)	0.71
Female home	.28*			.304	(.07)	.23*			.106	(.02)	0.01
Female work	.05			.021	(.01)	.01			−.036	(−.02)	0.05
Adjusted R^2		.22		.29			.23		.34		
Constant		1.13		−0.62			2.73		0.87		

NOTE: Unstandardized regression coefficients are given first with betas in parentheses.
* Individual coefficient significant at .05 level.
+ Differences in *r* or *b* between blacks and whites are significant at .05.

abortion requires focusing on the 1982 GSS, where the oversample of blacks gives sufficient cases for reliable estimates. As evident in Table 3, the inclusion of the four attitudes mentioned above and a further measure of feminism, support for the ERA, substantially increases the amount of variance explained in the abortion attitudes of both blacks (from an adjusted R^2 of .22 to .29) and whites (from an adjusted R^2 of .23 to .34). In fact, attitude toward premarital sexual relations becomes the best predictor of abortion attitudes for both blacks and whites.

In Table 3, we present the comparison by race of the impact of the demographic-religious variables suggested by Combs and Welch and our attitudinal variables. Reviewing the zero-order correlations (columns 1 and 6) shows seven variables which are differently associated with prochoice attitudes for blacks and for whites. The negative association of church attendance with prochoice attitudes is significantly stronger for whites than for blacks. Respondents of both races who have a sexually permissive attitude, are not fundamentalist, or are pro-ERA are more likely to be prochoice, but these are also significantly stronger relationships for whites than for blacks. The relationship of age with abortion attitudes (more prochoice if younger) and of region (current residence and region at age 16—more prochoice if from the North) is stronger for blacks than for whites. All seven variables continue to have somewhat different effects by race in the final regression, but only a few of these differences are statistically significant.

We test the significance of differences by race in the effects of our predictor variables by including race and a separate term for the interaction of race with each variable in an equation for the entire sample. This regression shows when effects are significantly different in magnitude or direction for blacks and whites. The *F*-tests for these interactions are given in column 11 of Table 3. We find only two significant differences: attitudes toward premarital sexual relations and age. The nature of these differences is most readily visible by examining the coefficients of regressions run separately by race. As we see in columns 4 and 9 of Table 3, which report these separate regressions, attitude toward premarital sexual relations, while significant for both groups, is considerably more important for whites. Age affects both blacks ($b = -.012$) and whites ($b = +.016$), but in opposite directions. For blacks, all other things being equal, younger respondents are more likely to be prochoice, while for whites, older respondents are more likely to be prochoice once factors such as fundamentalism and feminism are controlled.

Another type of difference evident in the paired regressions consists of those variables which are significant for one race but not for the other. If these differences were purely a function of sample size, we would expect all these variables to be significant only for the whites. Instead, we find a pattern of four variables which are significant only for whites and two variables which are significant only for blacks. Both measures of social class—income and education—are independently significant for whites but neither of them is for blacks. As would be expected, Catholicism is significant for only white respondents. Urbanness of present community is significant for whites but not for blacks, while region at age 16 is significant for blacks but not for whites. That is, whites presently living in rural areas are less prochoice, but for blacks, it is not the ruralness of their current community but growing up in the South that means they are less prochoice in abortion attitudes. We interpret this to mean that blacks tend to carry some normative aspects of their rural upbringing with them that influence their abortion attitudes even if they are presently living in the urban North, while whites who move to more rural or urban areas are more likely to conform to the norms prevalent there. Women are not more prochoice than men in the zero-order relationships, but a sex difference emerges once other factors are controlled. The positive effect of being female is strong enough to be significant for blacks but not for whites. Since this difference emerges even when other feminist attitudes are controlled, it is consistent with but not explained by the greater traditionality of black men that Ransford and Miller found in other sex role attitudes (1983).

DIFFERENCES IN ATTITUDE STRUCTURE

Having specified some differences in the determinants of abortion attitudes for blacks and whites, we turn now to differences in the structure of the attitude itself. Previous studies have typically either defined abortion attitudes in terms of the degree of total acceptance for all reasons (as we do, following Combs and Welch, above), or in terms of a dichotomy of the more generally accepted "hard" reasons of mother's health, rape, or birth defect in the fetus, and the less popularly supported "soft" reasons that the mother is single, can't afford, or wants no more children. Either of these strategies is useful for distinguishing the absolutists: those respondents who accept abortion for all six reasons (21 percent of all the blacks and 39 percent of all the whites) and those who oppose abortion for all six reasons (15 percent of all the blacks and 8 percent of all the whites). However, using either of these cutting points obscures the responses of those respondents who have a mixture of pro- and anti-abortion attitudes.

To clarify the structure of attitudes among the nonabsolutist blacks and whites, we employ a scale developed by Deitch (1982). Those who accept all the "hard" reasons but not all the "soft" reasons are classified as having mild anti-abortion attitudes (33 percent of the blacks and 39 percent of the whites). Those who oppose one or more of the "hard" reasons but accept at least one reason overall are classified as having strong anti-abortion attitudes (32 percent of the blacks and 15 percent of the whites). Table 4 presents the percentages accepting each of the individual reasons for those who have mild and strong anti-abortion attitudes.

Among respondents with a mild anti-abortion attitude, blacks and whites show a different pattern of acceptance of the "soft" reasons. Whites make few distinctions between the three "soft" reasons—their acceptance varies only 6 percent between "wants no more" (18 percent) and "can't afford" (24 percent). On the other hand, blacks make more distinctions among the social reasons they accept as legitimate—their acceptance varies 21 percent between the three reasons (only 13 percent accept abortion for a single woman but 34 percent want abortion to be legal for a woman who can't afford another child).

For whites with a strong anti-abortion attitude, there appear to be three distinct levels of acceptance of the individual reasons: woman's health, the other two "hard" reasons of rape and defect, and the three "soft" reasons which each receive less than 10 percent acceptance. Although blacks with a strong anti-abortion attitude share the major break in support between woman's health and the other two "hard" reasons, the distinction between the "hard" and "soft" reasons is less

Table 4. Percentage Distribution of Individual Reasons Accepted by Mild and Strong Anti-abortion Respondents by Race, 1982

	Mild Anti-abortion[a]				*Strong Anti-abortion*[a]			
	Black		*White*		*Black*		*White*	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Woman's health	100	na	100	na	88	12	83	17
Rape	100	na	100	na	44	56	48	52
Birth defect	100	na	100	na	37	63	38	62
Can't afford	34	66	24	76	22	78	9	91
Single	13	86	21	79	17	83	8	92
Wants no more	24	76	18	82	15	85	9	91
% of race in this category	33		39		32		15	
N	170		516		156		193	

[a] Mild anti-abortion attitude consists by definition of those who say "yes" to all three "hard" reasons (health, rape, defect) and "no" or "don't know" to one or more of the "soft" reasons. Strong anti-abortion attitude consist of those who say "no" to one or more of the "hard" reasons but did not answer "no" to all reasons.

pronounced. Blacks are more likely than whites to accept some of the social reasons even when they have a strong anti-abortion attitude. In addition, blacks again show a pattern of more differentiation between the three social reasons: whites vary in their acceptance between the three social reasons only by 1 percent, while blacks vary by 7 percent (15 percent want abortion legal for women who want no more children but 22 percent do for women who can't afford another child). It appears that among blacks, social reasons, especially economic need, are perceived as "harder" or more legitimate than whites see them. Moreover, the nonabsolutist blacks are consistently differentiating more between the three "soft" reasons they accept as legitimate than are whites.

Summary and Conclusions

All five hypotheses of this article are at least partially supported. First, new and better data on blacks have made it possible to see that the apparent convergence between black and white abortion attitudes in the period 1972–80 is misleading. The gap in prochoice attitude is as great in the 1982–84 period as it was 10 years before, and the pattern of lesser black support for legal abortion is consistent on an item-by-item basis over the decade.

Known differences in the characteristics of the black and white populations (in SES, religiosity, and Southern culture) explain some but not all of this persistent difference. Contrary to our hypothesis, the race gap that remains once these factors are controlled is not significant in 1983–84. The race difference is still significant in 1982. In those years where data are available, controlling directly for attitudes toward women and sexuality increases the explanatory value of the model but actually increases the racial gap between black and white abortion attitudes. These attitudinal variables are not available for the years of least difference. Nonetheless, it appears that the difference between the races is not solely attributable to their different social positions and attitudes.

Looking more closely at 1982, the pattern of determinants of abortion attitudes is found to differ for blacks and whites. Education and income, rural residence, and Catholicism are significant for whites but not blacks, while gender and Southern upbringing contribute significantly to the explanation of blacks' attitudes but not whites'. Attitude toward premarital sex is significant for both blacks and whites, but significantly stronger for whites, while the effects of age are equal and opposite in the two groups. Contrary to our fourth hypothesis, however, the link between abortion and specifically feminist attitudes is not significantly different for blacks than for whites.

For those respondents who differentiate their acceptance of legal abortion by situation, the pattern of prochoice attitudes differs by race as predicted. Black respondents are more likely to accept the social reason "married and can't afford more children" as legitimate than whites are, and more likely to accept this reason than that the woman is single and does not want to marry the father. White respondents who differentiate situationally are about as likely to accept being single as being poor as a legitimate reason and draw fewer distinctions among the social reasons in general.

One surprising difference between the two races is in the effect of age. With these demographic and attitudinal factors controlled, it is the older whites who are prochoice, while it is younger blacks who, all other things being equal, accept legal abortion. Since age is typically related to more conservative attitudes, the residual positive effect of age for whites is somewhat paradoxical. One possibility is that, among whites, there might be a small conservative component in support for legal abortion that emerges when the more obvious conservative factors (fundamentalism and antifeminism) are controlled. This conservative component may be an expression of symbolic racism (Sears and others, 1979), that is, seeing legal abortion as an economic alternative to supporting a supposedly black welfare population. The absence of such an antiwelfare component in black prochoice attitudes might explain both the differing effect of age and some of the lesser support for legal abortion among blacks. Another possibility is a cohort effect in which older whites, but not older blacks, have more experience with illegal abortion and resist return to the danger and social stigma of that period. But this speculation deserves empirical investigation that is beyond the scope of the present article.

Given the differences in black and white prochoice attitudes found here, exploring the political bases of these positions seems a promising direction for further research. Luker's insightful study of prochoice and prolife activists reveals how abortion attitudes are intricately interwoven with beliefs about motherhood, children, female roles, sexuality, and morality. Since each side is defending a world view that it does not and can not question, the heated ideological nature of the debate is understandable. The particular attitudes involved may vary by subculture. Some authors, including Combs and Welch, refer to the abortion-as-genocide theme as a possible explanation of black anti-abortion attitudes. We think this is possible, but that this is not the major explanation of subcultural differences.

Our findings suggest that blacks, like whites, connect legal abortion to other aspects of women's rights, but that the pattern of support varies. Attitudes toward premarital sex are not as important for blacks, and the situation of a single woman who does not want to marry the father is also perceived as a less acceptable reason for abortion than is economic need. Control over women's sexuality thus seems less central to blacks' attitudes toward abortion. We suggest the explanation lies in the definition of women and family values found in black culture. Black women are seen as crucial support for family and community, and women are considered primarily in terms of the home, but the home is not particularly defined by marriage. Family values are important, but families need not be male-headed to be legitimate.

In sum, white as well as black attitudes toward legal abortion seem to be based on cultural and political factors that are largely similar but include some important differences by race. After controlling for demographic, religious practice, and attitudinal variables, we are left with an irreducible difference between blacks and whites. Black anti-abortion attitudes, like white views, reflect traditional values of their subculture. For blacks this may include accepting single women as adequate mothers. For whites, there may be an element of symbolic racism. For both, the role of personal experience in the formation of prochoice attitudes remains to be investigated.

References

Arney, William, and William Trescher
1976 "Trends in attitudes toward abortion, 1972–1975." Family Planning Perspectives 8 (3):117–24.

Blake, Judith, and Jorge H. DelPinal

1980 "Predicting polar attitudes toward abortion in the United States." Pages 29–56 in James Burtchaell (ed.), Abortion Parley. Kansas City: Andrews and McNeel.

Conover, Pamela Johnston, and Virginia Gray

1983 Feminism and the New Right: Conflict over the American Family. New York: Praeger Publishers.

Combs, Michael, and Susan Welch

1982 "Blacks, whites, and attitudes toward abortion." Public Opinion Quarterly 46 (4):510–20.

Davis, James A., and Tom W. Smith

1982 General Social Surveys, 1972–1982: Cumulative Codebook (NORC edition). Roper Public Opinion Research Center, University of Connecticut, Storrs.

Deitch, Cynthia

1982 "The new right, feminism, and abortion: patterns of public opinion, 1972–1980." Paper presented at the American Sociological Association meetings.

Evers, Mark, and Jeanne McGee

1977 "The trend and pattern in attitudes towards abortion: 1965–1976." Paper presented at the American Sociological Association Meetings, Chicago.

Granberg, Donald

1978 "Prolife or reflection of conservative ideology? an analysis of opposition to legalized abortion." Sociology and Social Research 62 (3):414–29.

Granberg, Donald, and Beth Granberg

1980 "Abortion attitudes, 1965–1980: trends and determinants." Family Planning Perspective 12:250–61.

Lipson, Gerald, and Dianne Wolman

1972 "Polling Americans on birth control and population." Family Planning Perspectives 4:39–42.

Luker, Kristin

1984 Abortion and the Politics of Motherhood. Berkeley: University of California Press.

Mueller, Carol

1983 "In search of a constituency for the 'new religious right.' " Public Opinion Quarterly 47:213–29.

Pomeroy, Richard, and Lynn Landman

1972 "Public opinion trends: elective abortion and birth control services to teenagers." Family Planning Perspectives 4:44–55.

Ransford, H. Edward, and Jon Miller

1983 "Race, sex and feminist outlooks." American Sociological Review 48 (1):46–59.

Sears, David O., Carl P. Hensler, and Leslie K. Spier

1979 "Whites' opposition to 'busing': self-interest or symbolic politics." American Political Science Review 73 (2):369–84.

Smith, A. Wade

1984 "Evaluating the products of alternative sampling methods." Paper presented at the American Sociological Association Meetings, San Antonio.

Taylor, D. Garth

1980 "Procedures for evaluating trends in public opinion." Public Opinion Quarterly 44:86–100.

Westoff, Charles, Emily Moore, and Norman Ryder

1969 "The structure of attitudes toward abortion." The Milbank Memorial Fund Quarterly 47:11–36.

Picking the Winners: Politician vs. Voter Predictions of Two Controversial Ballot Measures

JAMES B. LEMERT

PUBLIC opinion researchers have known for some time that members of the public often can be wildly incorrect when estimating how other members of the public feel about issues. This general finding has been called *pluralistic ignorance* by some (e.g., O'Gorman and Garry, 1976). Roughly at the same time, students of legislative representation such as Miller and Stokes (1963) were developing evidence—disputed as an artifact in some quarters—implying that legislators' perceptions of their constituents' opinions were more accurate as issue salience/controversy increased.[1]

Justifications of representative democracy implicitly assume that politicians are actually capable of perceiving the policy wishes of their constituents. Obviously, any model of representation that requires

---

[1] See especially Achen (1977), who asserts that use of the correlation coefficient inevitably meant that members of the 1958 House of Representatives would show greater "accuracy" in their estimates of constituency opinion on civil rights, where between-district variation was greatest, than on social welfare and foreign policy, where regional differences were not as great. In the present study, both measures studied were controversial and both displayed marked differences in support between areas of the state. However, unlike Miller-Stokes and similar "representation" studies, we asked for predictions of *statewide* results, not of results in any district(s).

Abstract Both politicians and voters were asked to predict outcomes of two Oregon ballot measures in 1982. As expected, politicians' predictions always were closer to the mark than voters' were. Further, voters showed stronger signs of wishful thinking (the "Looking-Glass effect") in their predictions than did politicians. Using published preelection polls apparently improved politicians' accuracy in 1982, as well as voters' accuracy in a separate 1984 survey. No other sources of data improved predictive accuracy. Findings have implications for theories of representative government and are consistent with a new theory of public opinion.

James Lemert is Professor and Director of the Division of Communication Research, School of Journalism, University of Oregon.

Public Opinion Quarterly Vol. 50:208–221 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-208/$2.50

elected politicians to consider public opinion in decisions assumes that these decision makers' perceptions of public opinion are accurate "enough." But what is "enough"? Though Hedlund and Friesema found that, in each case, Iowa legislators were right much more often than wrong in predicting how their constituents would vote on four ballot measures, these authors interpreted their results as raising serious questions about representative democracy (1972: 410–11).

A major problem with Hedlund and Friesema's study design is that it considered only predictions made by politicians, merely comparing various subsets of them against other subsets. This seems a tenuous basis on which to conclude that the entire theory of representative democracy is threatened. In the absence of a baseline that is external to the sample of elected politicians, it is difficult to say what percentage of accuracy is sufficient. Fortunately, however, the pluralistic ignorance literature suggests—almost by definition—that politicians should at least do better than the mass public.

Lemert (1981: 12) has based a key part of his public opinion theory on the as-yet-untested proposition that perceptions of public opinion held by politicians and other decision makers will be more accurate than perceptions held by ordinary citizens, as a result of a series of constraints experienced by politicians. Constraints may include such things as (1) questions from journalists about decision makers' public opinion perceptions and their evidence for them, (2) contrary perceptions offered by rival decision makers, (3) uncertainty as to whether polls or other independent sources of public opinion data might surface at some future time, and (4) at least for elected politicians and their advisors, the continuing opportunity (or threat) presented by the need to face the voters.

No such pressures are felt by ordinary members of the mass public (Lemert 1981: 13–15). Without these constraints, the result for ordinary members of the public can be what Fields and Schuman (1976: 445) call the Looking-Glass perception: "In the absence of strong counter-forces, a large proportion of people feel that the world they live in agrees with their own opinion on public issues."

A kind of pressure upon ordinary citizens has been proposed by Noelle-Neumann in her writings about the "Spiral of Silence." In her book she tends to assume a kind of "quasi-statistical" accuracy on the part of ordinary citizens, but even there, she notes, inaccuracy can be created through the Looking-Glass effect and through inaccurate media portrayals of dominant opinion (1984: 124–27; 157–64). And in an earlier journal article (1974: 45), she explicitly asserts that the spiral of silence itself can mislead both the vocal, active people and the silent ones about the true distribution of opinion.

In the present study, we treat the way people voted in an election as an indicator of their "true" attitude on that issue. (Their reported vote is obtained *before the outcome of the election is known.*) We then make two predictions: (1) politicians will be much more accurate predictors of outcomes for two highly controversial state ballot measures than will a comparison group of voters, and (2) whatever apparent signs of the Looking-Glass phenomenon occur will be stronger among voters than among politicians.

This study cannot directly test Lemert's argument that greater pressure or constraint makes politicians more accurate. Nevertheless, his argument predicts considerably higher accuracy among the politicians. If this prediction is fulfilled, then we can begin to consider alternative explanations for it.

Method

For at least two reasons, asking voters to predict the outcome of elections may be a fairer test of pluralistic ignorance than is the practice of asking a general survey sample to estimate survey responses given by others. First, general samples obviously can include respondents who lack acquaintance with the issue(s) being asked about, thus producing artifactual overestimates of the breadth and depth of pluralistic ignorance. In contrast, people who have just voted on an issue presumably have some acquaintance with it. Second, issues in elections have a kind of history and concreteness that issues asked about in surveys may not. For example, if one is in the habit of trying to predict election results, one finds out the actual result soon enough. But if one is trying to predict attitudes toward racial integration in Detroit, for example (Fields and Schuman, 1976), how does one find out whether one is right or wrong?

The principal comparisons will be made between politician and voter predictions of two ballot measure outcomes in November 1982. In addition, certain supplementary data will be used from a second exit poll of voters, conducted this time at the November 1984 election.

The 1982 election in Oregon featured two highly publicized and controversial state ballot measures: Measure 3 (a property tax limitation measure patterned after California's Proposition 13) and Measure 6, which would have eliminated statewide land-use planning controls. The 1984 election included another property tax limitation measure which was virtually identical to the 1982 proposal. The three measures had characteristics we thought were ideal for our purposes. First, there was considerable uncertainty about whether each would pass, because polls early in the campaigns had shown all three of them with large leads while more recent polls showed their leads eroding rapidly. Second, each issue had a considerable track record in Oregon, close variants of each having been defeated in the past. Third, each issue clearly was salient to significant numbers of voters, judging by how frequently and rapidly petitioners had placed each on the ballot.

THE VOTER SAMPLES

The 1982 and 1984 voter groups were produced by a multistage, stratified random sampling procedure. The first stage was to obtain a sample of Eugene-Springfield (OR) metropolitan area precincts (29 in 1982, 41 in 1984), stratified by relative wealth and past liberal-conservative voting records. Precincts were chosen randomly within each of these strata. The next stage involved a sample of voting times. On the basis of Lane County Election Department estimates of turnout by daypart, half the interviewers went to precincts between 5 P.M. and the closing of the polls at 8 P.M., a quarter went between 1 P.M. and 5 P.M., and a quarter between 8 A.M. and 1 P.M. Within this overall time allocation, dayparts were assigned randomly to precincts. Finally, each interviewer was given a procedure for selecting specific interview respondents. A "skip interval" number between one and five was randomly assigned to each precinct. Interviewers given the interval number three, for example, would have to wait for three voters to leave before interviewing the next person. Interviewers resumed counting after completing each interview. Each interviewer was to obtain 10 respondents during the assigned time period. The 1982 survey produced 283 completed interviews in 29 precincts; the 1984 survey produced 408 in 41.

THE SAMPLE OF POLITICIANS

Mail questionnaires were sent before the 1982 election to 115 prominent Oregon politicians, all but three of whom were themselves involved in a race for state or U.S. office. On the assumption that pressure toward accuracy would be greater when a candidate had an opponent, only candidates with opponents were sampled. Three prominent elected state executives who were not directly involved in the 1982 election also were included in the mail sample. Their actions and decisions often had been intimately associated with land-use planning and property tax issues, so they could be expected to be constrained toward accurate perceptions of public opinion.

Names and addresses of candidates were obtained from the Oregon Secretary of State's office. Assuming normal delivery of the mail, each questionnaire arrived at the listed address seven days before the election. All politicians were instructed to mail their responses back so that the envelope would be postmarked *before* election day. Of the 115 questionnaires sent out, 58 usable replies were received. Four replies were discarded because they were postmarked too late, two envelopes were returned as undeliverable, and three politicians sent notes saying their questionnaire was received too late. The effective completion rate, therefore, was 58 of 106, or 54.7 percent.

The most common office sought by respondents was in the State House of Representatives (35 replies), with the next most common (13 replies) being the five U.S. Congress seats plus three contested state executive offices. Slightly more than half the replies came from incumbents (or, in a few cases, their campaign managers).

THE PREDICTION QUESTIONS

Similar questions were asked voters (1984) and both voters and politicians (1982) about each ballot measure. Before any other questions, respondents were first asked to predict whether each measure would pass or fail. Then they were asked to estimate the percentage of "yes" votes the measure would get from Oregon voters.[2]

Results

Both the 1982 and 1984 property tax measures were defeated narrowly by Oregon voters. The 1982 measure received 49.5 percent of the

---

[2] Question wordings differed slightly because voters were questioned in person while politicians filled in a mail questionnaire. Here is how politicians were asked the questions:

> Please write in your best estimate, based on the information you presently have, about how Measure 3—the property tax limitation measure—will do, *statewide, in percentage terms*.
>
> I think Measure 3 will get ____ percent "*yes*" votes statewide and thus will (circle one) pass. fail.

The question for Measure 6 was identical except for the measure's number and the substitution of "the LCDC 'repeal' " for "the property tax limitation measure."

Interviewers asked the comparable questions of voters as follows:

> I'd like to ask you to try your hand at predicting whether Oregon voters will pass Measures 3 and 6. . . . First, let's do Measure 3, the property tax limitation measure. Do you think Oregon voters will approve Measure 3, or do you think it will fail, statewide? What percentage of the vote will the YES side get, then, given that you think it will . . . (pass) (fail)?

The questions for Measure 6 were similar.

In 1984, the questions to voters differed slightly from the 1982 voter questions by explicitly asking them to ignore their own wishes:

> As you know, there's been lots of predicting about how Measure 2, the property tax limitation measure, will do. I'd like you to take a stab at guessing whether Oregon voters will pass Measure 2. Regardless of how you yourself voted on it, do you think Measure 2 will pass, or will it fail? What percentage of the vote will the YES side get, given that you think Measure 2 will . . . (pass) (fail)?

In all interviews with voters, interviewers were instructed to try to get an exact percentage estimate, rather than a range of percentages.

Table 1. Voters' vs. Politicians' Predictions of Two State Ballot Measures' Outcomes

	Measure 3			*Measure 6*		
	Predict Will Pass	*Predict Won't Pass*		*Predict Will Pass*	*Predict Won't Pass*	
Sample	%	%	*(N)*	%	%	*N*
Voters	36.4	63.6	(275)[a]	38.6	61.4	272[a]
Politicians	26.8	73.2	(56)[a]	10.5	89.5	57[a]
Total	34.7	65.3	(331)	32.6	67.4	(329)
	χ^2 = 1.484, Ns, with 1 df			χ^2 = 15.385, $p < .001$, with 1 df		

	Estimated % "Yes" Votes			
	Measure 3		*Measure 6*	
	%		%	
Actual	49.5		44.9	
Voters' mean estimate	45.8*	N = 277	46.7	N = 271
Politician's mean estimate	48.1*	N = 56	45.2	N = 56

	Variance in Estimates	
Voters	162.8686**	155.5001**
Politicians	10.0299**	20.3852**

* Means differ at $p < .05$ (t = 2.54, using separate error estimates and reduced degrees of freedom).

** On each measure, voters' variance significantly ($p < .001$) larger than politicians' by F-max test.

[a] Numbers of cases will vary from table to table because of incompletes.

vote, carrying most counties by a narrow margin. The 1984 result was very similar, the tax limitation measure receiving 49.3 percent of the vote and once again carrying most counties in the state. In contrast, the 1982 land-use measure got only 44.9 percent statewide and was defeated in most counties. Unless specifically noted, all comparisons reported here are from the 1982 mail and exit surveys.

Table 1 provides several tests comparing the accuracy of voters and politicians on each ballot measure. Majorities of both groups said each measure would fail, and all their estimated percentages had means of less than 50 percent. Nevertheless, all between-group comparisons favored the politicians: In all four cases, politicians' predictions were closer to the outcome than were voters'. For the property tax measure, politicians were only marginally more likely than voters to think the measure would fail, but their mean percentage estimate was significantly closer to the outcome (p = .011). As for the land-use measure, politicians were much more likely than voters to predict its defeat. The mean estimate for politicians was only 0.3 percentage points above the actual outcome (44.9 percent), while voters were 1.8 percentage points too high. Over half the politicians' estimates came within 3 percentage points of the actual result for the land-use measure, compared to less than one in five voters ($p < .001$ by chi-square test).

Variances in voters' percentage estimates were much higher than for estimates by politicians, requiring both *t*-test comparisons in Table 1 to assume heterogeneity of variance. Voter estimates ranged all the way from 2 to 91 percent.

This much greater range among voters' estimates suggests another possible explanation for politicians' greater accuracy: Perhaps accuracy differences reflect politicians' greater knowledge and sophistication, rather than constraint. Most sophisticated observers of U.S. elections know that almost all outcomes will fall between 40 and 60 percent. In this study, zero and 3.6 percent of politicians fell outside that range, compared with more than 30 percent (tax) and more than 25 percent (land use) of the voters.

Judging by a succession of tests, however, it won't be easy to explain voter-politician differences by means of the sophistication argument. First, even if we delete from comparisons everyone whose estimates fell outside the 40–60 percent range, voters' mean estimates hardly changed at all. A second variant of the sophistication argument notes that voters tended to use "round numbers" (i.e., multiples of five) in their estimates, while politicians used numbers in between, thus presumably putting politicians' estimates closer to the real-number outcome of 49 percent for the tax measure. Even after we delete both all cases outside the 40–60 percent range and all estimates made in multiples of five, accuracy differences remained robust. As with all the previous tests of percentage estimates, politicians' estimates clustered much more closely around the actual results.

What about the Looking-Glass effect? Were there any apparent signs of it, and if so, did the effect seem stronger among voters than politicians? The answers seem to be yes—and yes. We asked voters to tell us how they voted after they had made their predictions, but we did not ask politicians directly about their own attitudes toward the two ballot measures. For politicians, the best indicator we have of their own attitudes was their affiliation with either the Democratic or Republican parties. Generally speaking, both Oregon ballot measures were strongly supported by conservative groups whose membership overlapped with major elements of the Oregon Republican Party, and were strongly opposed by almost all Democratic politicians in the state. (Many moderate Republicans also opposed both measures, however.)

Let us consider Looking-Glass results for voters first, for politicians second, and then compare the apparent strength of the effect between the two groups. Table 2 presents all these results.

Among voters, how they voted very strongly predicted a belief that

Table 2. Apparent Looking-Glass Effects Among Voters, Politicians

	Measure 3 (Tax)			Measure 6 (Land Use)		
	Predict Will Pass	Predict Won't Pass		Predict Will Pass	Predict Won't Pass	
	%	%	(N)	%	%	(N)
Subsample						
Voted "yes" on this measure	70	30	(60)	74	26	(82)
Voted "no" on this measure	25	75	(194)	19	81	(189)
Rep. pol.	48	52	(23)	17	83	(24)
Dem. pol.	9	91	(32)	6	94	(32)

χ^2 Comparisons	Measure 3	Measure 6
"Yes" voters vs. "no" voters	$\chi^2 = 39.07$, $p < .001$, with 1 df	$\chi^2 = 65.11$, $p < .001$, with 1 df
Rep. vs. Dem. pol.	$\chi^2 = 8.49$, $p < .01$, with 1 df	$p < .20$ by Fisher's Exact Probability Test
"Yes" voters vs. Rep. pol.	$\chi^2 = 2.65$, $p < .15$, with 1 df	$\chi^2 = 23.71$, $p < .001$, with 1 df
"No" voters vs. Dem. pol.	$\chi^2 = 2.89$, $p < .10$, with 1 df	$\chi^2 = 2.84$, $p < .10$, with 1 df

	Mean Estimated Percentage "Yes"			
	Measure 3		Measure 6	
Subsample	%		%	
Voted "yes" on measure	51.15^{a}	($N = 61$)	$53.96^{a,c}$	($N = 83$)
Voted "no" on measure	$44.16^{a,b}$	($N = 194$)	43.27^{a}	($N = 167$)
Rep. pol.	49.00^{d}	($N = 24$)	$46.33^{c,e}$	($N = 24$)
Dem. pol.	$47.26^{b,d}$	($N = 31$)	44.26^{e}	($N = 31$)

[a] Means differ at $p < .001$, based on *t*-test using pooled error estimate.

[b] Means differ at $p < .01$, based on *t*-test using separate error estimates and reduced degrees of freedom.

[c] Means differ at $p < .001$ based on *t*-test using separate error estimates and reduced degrees of freedom.

[d] Means differ at $p < .05$, based on *t*-test using pooled error estimate.

[e] Means differ at $p < .10$, based on *t*-test using pooled error estimate.

their own side would win, regardless of whether their own vote was "yes" or "no" and regardless of which measure it was. Looking across both measures, at least 70 percent of those who voted "yes" also thought their measure would win, while at least 75 percent of "no" voters thought that measure would be defeated ($p < .001$ for each ballot measure). Similar findings occurred for estimates of percentages. "Yes" voters estimated "yes" votes at a mean 7 percent higher (tax measure) and 10 percent higher (land use) than the means for "no" voters ($p < .001$ in each case).[3]

As for the politicians, the evidence for a Looking-Glass effect was stronger for the tax relief measure. Republicans were much more likely to think that the tax limitation would win ($p < .01$), and their mean percentage estimate on the measure also was higher than the Democrats' ($p < .05$). Results were in the same direction for the land-use measure, but not significantly so.

Already the Looking-Glass evidence among politicians seems weaker than among voters. Is there a tougher and more direct way, though, to compare the two? One highly plausible way would be to compare those whose attitudes ought to be distorting their perceptions in the same directions. If, for example, Democratic politicians are closer to the actual results than are "no" voters, we could argue that Democratic politicians' Looking-Glass effect was the weaker of the two. Then we would do a similar comparison between Republican politicians and "yes" voters. Table 2 again contains the results bearing on these comparisons. Looking first at the question of whether the measure will win or lose, we find that in all four cases ("yes" voters vs. Republican politicians and "no" voters vs. Democratic politicians, on each ballot measure) the politicians less often predicted the measure would win. Looking now at the bottom of Table 2, we find that the politicians are closer to the actual percentage outcome in all four cases. Summing all these comparisons across both percentage estimates and "win-lose" questions, in all eight cases politicians' predictions were closer to the mark, and five of these eight differences reached at least the $p < .10$ level.

A slightly different way of comparing the strengths of apparent Looking-Glass effects would be to look at the *distances* of estimates from actual two-digit percentage outcomes, once again comparing Republican politicians with "yes" voters and Democrats against "no" voters. For each 1982 ballot measure, exactly half the Republican politicians made estimates that were 3 percentage points or less away from the actual results, compared to less than one in four (taxes) and one in nine (land-use planning) "yes" voters. These differences were significant at $p < .05$ and $p < .001$, respectively, by chi-square tests. Nearly three-fourths of the Democratic politicians were within 3 percentage points of the tax outcome, compared to less than 18 percent of the "no" voters ($p < .001$). As for the land-use outcome, more than 61 percent of the Democratic politicians were within 3 percentage points of the outcome, while only 27 percent of the "no" voters were ($p < .001$).

---

[3] The same apparent Looking-Glass effect was present among 1984 voters—and at least as strongly. Some 74 percent of the "yes" voters on the property tax measure thought it would pass, while 85 percent of the "no" voters thought it would be defeated. In addition, "yes" voters estimated "yes" votes at a mean that was 53.8 percent while "no" voters' mean estimate was only 41.2 percent (all differences significant at $p < .001$, by χ^2 and t-tests, respectively).

If politicians were better predictors than the voters, was it because politicians had more public opinion data to aid them? The top half of Table 3 compares the numbers and types of data sources used by politicians (in 1982) and voters (in 1984). In each survey, respondents could cite as many as 10 possible sources of public opinion data. Though the two lists of 10 sources weren't completely identical,[4] comparison of the items they did have in common (e.g., published polls, media reports and letters to the editor) clearly shows that politicians used them much more often than did voters (Table 3). Overall, politicians used a mean of 4.9 of their 10 sources, while voters used only 2.2 of their 10.

Apparently, then, politicians used a greater variety of sources to make their predictions than did voters. It is tempting to test whether the greater variety of sources used by politicians (in 1982) made them better predictors than voters were in 1984. But when the question is put this way, the hazards of assuming complete comparability are obvious. Not only were the lists of sources not identical for the two elections but, in 1982, the sources were for *two* ballot measures while in 1984 they were for only one. Nevertheless, we can stay within each data set and use sources to predict accuracy differences *among* politicians in 1982 and *among* voters in 1984.

When we look separately at each one of the 10 different sources, only one—published polls[5]—predicted accuracy differences among politicians as well as among voters (see bottom half of Table 3). For

---

[4] It was inevitable that the two lists of sources would not be identical because there was no real equivalent, for voters, of such obvious and important sources for politicians as contacts with constituents or impressions about voters from political colleagues. In order to try to be fair to voters, however, we separated such similar sources as "conversations with family and friends" from "conversations with acquaintances," so that one basic type of source—conversations—actually could produce a count of two sources for voters.

[5] The state's most prominent pollster, J. Roy Bardsley, told the author that his final published polls showed all three measures getting a plurality of "no" votes. However, in each case the number of undecided voters was larger than the margin of apparent defeat. Further, the time lags between final polling and election days ranged from one to two weeks. Final Bardsley and Haslacher, Inc. figures for 1982: Measure 3—37 percent yes, 45 percent no, with 18 percent undecided; Measure 6—38 percent yes, 40 percent no, with 22 percent undecided. The firm's final figures for the 1984 tax measure were 46 percent yes, 49 percent no, with 5 percent undecided. According to Bardsley, his published poll was the only one to predict the 1984 defeat of the tax measure.

Table 3. Sources of Data for Predictions

Politicians' Sources 1982	*% Using (N = 58)*	*Voters' Sources 1984*	*% Using (N = 408)*
Personal contacts, conversations	89.7	Family, friends	58.3
Published polls	77.6	Polls in news	15.4
Media reports	69.0	News reports	40.7
Contacts with groups	60.3	Meetings	4.2
Informants, colleagues	58.6	"Experts"	12.5
Editorials	37.9	Editorials and advertisements	30.4
Letters to the editor	29.3	Letters to the editor	12.5
Constituent letters & calls	27.6	Acquaintances	15.9
Private polls	22.4	Bumper stickers, lawn signs	11.8
Other	12.0	Other	13.2
Mean sources named: 4.862		Mean sources named: 2.167	

Sources as Predictors of Accuracy

Will 1982 tax limit pass?	Yes	No	Will 1984 tax limit pass?	Yes	No
Used polls[a]	24.4%	75.6%	Used polls[a]	29.8[b]	70.2[b]
Did not use polls[a]	30.8	69.2	Did not use polls[a]	41.0[b]	59.0[b]
Will measure 6 pass?	Yes	No	% estimate (means)	Tax Limit	
Used polls	4.4%[c]	95.6%[c]	Used polls	49.50%	"Yes"[d]
Did not use	30.8[c]	69.2[c]	Did not use	43.77%	"Yes"[d]
% estimate (means)	Tax Limit	Land Use	Used bumper stickers, lawn signs	40.55%	"Yes"[f]
Used polls	47.95%	44.74%[e]	Did not use	45.21%	"Yes"[f]
Did not use	48.25%	46.67%[e]			

[a] "Polls" always means published polls; 1982 respondents: politicians; in 1984: voters.

[b] By χ^2 test, $p < .10$.

[c] Fisher's Exact Test, $p = .0106$, two-tailed.

[d] Means differ at $p < .01$, $t = 3.05$.

[e] Means differ at $p < .001$, $t = 3.55$.

[f] Means differ at $p < .05$, $t = 2.20$.

politicians, using polls seemed to work only in predicting the land-use measure. With regard to the 1982 tax limitation measure, politicians who used the polls were in fact very slightly farther away than their colleagues from the actual percentage. Interestingly, six politicians carried their tax measure estimates to three decimal places and every one of them very precisely bracketed the actual outcome of 49.5 percent. None of the 10 sources—not public polls, not private polls, not any—differentiated these six politicians from their colleagues.

Turning to the voter sample from 1984, we find that relatively few voters cited published polls. But those who did were somewhat ($p < .10$) more likely to think the tax measure would fail, while at the same time their higher percentage estimate (a mean of 49.5) was significantly closer to the actual outcome (49.3 percent).

Of all sources used by politicians and voters, only one other source besides polls was significantly related to any measure of accuracy for either group. In this single case, though, this source (lawn signs/bumper stickers) was related to *greater error* among voters.

For the voters, relying on bumper stickers/lawn signs as a source might help "cancel" the benefits of following the polls, when both kinds of sources were used. This pattern of counteracting sources might help explain why more sources of information were not necessarily better, in terms of predictive accuracy. The total numbers of sources used did *not* predict accuracy differences among either the voters or the politicians.

Discussion

This study is the first directly comparing the accuracy of election predictions by politicians with those by members of the mass voting public. As predicted, politicians were better than voters at predicting the outcomes of two controversial ballot measures, and they displayed fewer and weaker signs of the Looking-Glass effect. By showing consistently lower variances among politicians than voters in their percentage estimates, the findings also suggest that it may be more important, first, to consider the pressures—or attributes—politicians have in common, instead of first trying to differentiate among them. The findings suggest also that it is premature to discard models of representative democracy that presuppose politicians' ability to make accurate judgments about constituent opinion.

To show that political decision makers' perceptions of public opinion are more accurate is to support Lemert's argument that decision makers' perceptions are in enough touch with political "reality" that their perceptions will change as the information coming to them about citizen attitudes changes. This proposition thus allows Lemert to argue that the mass media more often change "public opinion" through changes in the information reaching decision makers about citizen attitudes than they do through attitude change itself (see especially Lemert, 1981: 32–37, 213–217).

Why would politicians be less susceptible than voters to wishful thinking? Perhaps, instead of being constrained to avoid wishful thinking, politicians are simply more interested in, and savvy about, elections. Another way of putting this is that we might be comparing full-time professionals against amateurs in the prediction game. Interest alone does *not* seem to provide the explanation. Politicians' personal greater interest predicts their using a greater variety of indicators than voters for their predictions. This appears to have happened. But the total number of different indicators used bore *no* relationship to predictive accuracy within either the voter or the politician data sets. It may be that only some of these indicators enhance predictive accuracy, while others interfere with it.

It was true that using published preelection polls predicted greater accuracy, both among voters and among politicians. And it also was true that politicians were more likely than voters to use published polls. But if some other indicators can interfere with accuracy (bumper stickers/lawn signs for voters), it is hard to see how greater interest alone could lead *both* to using polls and *ignoring* certain of these other indicators. Perhaps interest needs to be accompanied by experience and savvy, both of which could lead to a more judicious use and interpretation of the available evidence.

We can't resolve here whether constraint or savvy—or both—explains why politicians were better able to resist wishful thinking. One reason we can't settle this issue is that greater constraint implies greater pressure toward accuracy, which in turn implies both interest in, and the seeking out of, preelection data. Further, a more judicious, savvy use of these data may itself be a constraint or "survival" characteristic for a politician (e.g., Converse et al., 1966:24; Hedlund and Friesema, 1972:406–7). One of many ways to measure constraint as an independent variable would be to contrast the accuracy of politicians' predictions by the salience of issues: The more salient the issue, the more accurate should be the prediction by politicians. In contrast, the more salient the issue to voters, the more likely that the wishful thinking/Looking-Glass effect might be.

Another issue to be considered here is whether the apparently weaker Looking-Glass effect among politicians could be an artifact of a less precise measure of politicians' attitudes toward the two measures. Suppose that votes cast for and against each measure reflected voter attitudes very precisely, but party affiliation much less precisely categorized politicians' own attitudes. Presumably, then, the relationship between own attitudes and predictions would appear to be weaker among politicians than among voters. It was indeed true that some moderate Republican politicians did oppose one or both of the two measures, and we had no way of identifying them because the returned questionnaires were anonymous. Therefore, these moderate Republicans were combined with more conservative Republicans, most of whom *did* support both measures, thus weakening any signs of wishful thinking among Republican politicians. While plausible, this argument simply cannot explain why Democrats—almost universally opposed to both measures—*also* showed a smaller Looking-Glass effect than did their comparison group of "no" voters.

Finally, let us consider one last threat to these results. The property tax measure would have been adopted as part of the Oregon Constitution had Eugene-Springfield voters not delivered a resounding margin of "no" votes that just barely offset the rest of the state. Could the fact that most of our sample of Eugene-Springfield voters were opposed to both measures have somehow produced the voter errors reported here? No, not at all. In fact, since both measures lost, whatever "bias" existed in our sample usually stacked the deck in the direction of *minimizing* voter error, not enlarging it. If we had weighted our sample of "yes" voters to reflect the overall state outcomes on each measure, we would have greatly *enlarged* voters' predictive errors in three of the four possible cases. Of the only six people to carry out their 1982 estimates to three decimals, every one of them did it only for the tax measure—and every one of them hit the actual outcome, almost to the fourth digit. That all six of them were politicians cannot be explained by either weighting or not weighting a sample of voters.

References

Achen, Christopher H.
1977 "Measuring representation: perils of the correlation coefficient." American Journal of Political Science 21: 805–15.
Converse, Philip E., Aage R. Clausen, and Warren E. Miller
1966 "Electoral myth and reality: the 1964 election." In Edward C. Dreyer and Walter A. Rosenbaum (ed.s), Political Opinion and Electoral Behavior: Essays and Studies. Belmont, CA: Wadsworth.
Fields, James M., and Howard Schuman
1976 "Public beliefs about beliefs of the public." Public Opinion Quarterly 40: 427–49.
Hedlund, Ronald D., and H. Paul Friesema
1972 "Representatives' perceptions of constituency opinion." Journal of Politics 34: 730–52.
Lemert, James B.
1981 Does Mass Communication Change Public Opinion After all? A New Approach to Effects Analysis. Chicago: Nelson-Hall.
Miller, Warren E., and Donald E. Stokes
1963 "Constituency Influence in Congress." American Political Science Review 67: 45–56.
Noelle-Neumann, Elisabeth
1974 "The spiral of silence: a theory of public opinion." Journal of Communication 24: 43–51.
1984 The Spiral of Silence: Public Opinion—Our Social Skin. Chicago: The University of Chicago Press.
O'Gorman, Hubert, and Stephen L. Garry
1976 "Pluralistic ignorance—a replication and extension." Public Opinion Quarterly 40: 449–58.

Election Predictions: An Empirical Assessment

WILLIAM BUCHANAN

POLLSTERS have been publishing predictions of national election results for nearly 50 years. There are now enough instances to make possible a comparison of the *empirical* with the *theoretical* standards of accuracy.

Accuracy of election predictions may be expressed by several measures, each useful for a different purpose. These were developed by the Social Science Research Council team which appraised the performance of the American pollsters after their 1948 fiasco (Mosteller, 1949: 54–61). Among these measures are the following:

1. The proportion of polls which correctly predicted the winner.

2. "Average error," the mean of the differences between the predicted percentage and the actual vote percentage, ignoring the sign. Most easily understood by newspaper readers, this average has the defect of concealing biases—all the prediction errors might be in favor of the more conservative party, or of the ultimate winner or of some other category.

3. Systematic error, the mean of the differences between the predictions and the election percentage, taking account of sign. This reveals the biases, but understates the size of the errors, since large overestimates cancel out equally large underestimates.

4. Nonsystematic error, the standard deviation of the individual er-

---

Abstract Examination of 155 poll forecasts in 68 national elections since 1949 shows that errors average nearly twice what statistical theory would indicate. Polls predict the division of vote between major parties better than individual party percentages, leading to 85 percent success in picking the winner. The worst failures occurred in a few elections where most polls went wrong. Liberal party votes are correctly forecast, conservatives slightly underestimated. Improved polling methods have not led to better forecasts.

William Buchanan is Professor of Politics at Washington and Lee University.

Public Opinion Quarterly Vol. 50:222–227 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-222/$2.50

rors around the mean of the errors in a series of elections. It tells how widely the predictions vary, discounting systematic error. It is not too meaningful to lay readers, but useful to statisticians since it permits one to compare the performance of polls with the deviations that would be expected for given sample sizes. It is these probability calculations that underlie the familiar "margin of error" statements that appear in all responsible reportage of poll forecasts. However, these calculations are based on probability theory, and do not take account of the actual performance of the polls.

These error calculations may be made for each party's vote as a percentage of the total vote, or as a percentage of the "major party" vote. The latter is what most consumers of polls are interested in, since it has political relevance. It predicts who will come in first and how far ahead of the second-place party.

The SSRC's report showed that through 1948 the polls' nonsystematic error was not excessive, but there had been a decided pro-Republican bias, as reflected by the systematic error measure.

Robert M. Worcester (1983) has recently edited a history of political polling that brings together actual and predicted results of 96 election polls in Australia, France, Germany, Britain, Japan, and the Netherlands. I have updated and supplemented these with Canadian, New Zealand and Australian, and U.S. predictions, bringing to 155 the total of forecasts made in 68 elections since 1949. These were the final published percentage predictions of the popular vote for a candidate, party, or coalition in national elections. The data set includes a large majority of such predictions made during the period. Sample size is not always given, but where it is, 1500 or 2000 is the norm, with occasionally larger and rarely smaller samples, none under 900. The assumption for purposes of analysis that the average sample size is 1500 is conservative.

Using the categories established in the SSRC report, average, systematic, and nonsystematic errors were calculated for the party that won the election and also for the one that came in second, each as a percentage of the total vote. Then the percentages of winner and runner-up were added and the winner's percentage of the two-party vote compared to the prediction. This leaves out a residue of votes for minor parties, plus, in a few instances where it was so published, an "undecided" percentage, these being of little interest to consumers of predictions. Since there were several competing forecasts in most elections it seemed sensible to subtract the predicted percentage from the actual percentage; thus errors with a *positive* sign in the calculations that follow are *under*estimates and those with a minus are overestimates of what that party actually got.

The published "margin of error" percentages are derived from the statistical formula for the standard error:

$$1.96\left(\sqrt{\frac{pq}{n}}\right)$$

where p is the proportion we are trying to estimate, $q = 1 - p$, and n is the number of cases in the sample.

The statistical analog for this calculation is a vat containing millions of marbles (equivalent to the electorate) of which about half are white (the leading party's vote). The pollster scoops up a bucketful of marbles (roughly 1500) time after time, counts them, throws them back, stirs the vat and scoops up another bucketful (for 155 samples). If done with marbles we could count on close to 95 percent of the sample estimates varying from the proportion of white marbles in the vat by no more than than the margin of error based on the formula above, in this case 2.5 percent: $1.96\sqrt{(.5 \times .5) / 1500} = .025$.

Merely describing this ideal sampling arrangement is enough to show the contrast to what election forecasters are really doing. Voters aren't susceptible to being stirred and replaced; similar ones are clumped together and surveys sample these neighborhood or telephone prefix clusters. Some aren't home and some don't have phones. Some haven't made up their minds and some won't vote, though they may think they will (we don't know what color marbles these are, or indeed whether they are marbles at all). Each poll has its own adjustment procedure, often a valued trade secret, for coping with these problems. The customary "margin-of-error" stipulations in the media are based upon the formula for simple random sampling, which rounds off to plus-or-minus 2.5 percent for 1500 cases, 2 percent for 2000 cases. With adjustment for cluster sampling these figures would be around 3.5 percent and 3 percent, respectively.

The 95 percent probability criterion calls for an *average* of one "bad" prediction—i.e., one outside the specified margin—in every 20 attempts. National elections are infrequent and we need the full sample of 155 to make use of the principle, which would call about 8 errors of this size.

The calculations of the relevant averages appear in Table 1. The average absolute error for the 155 predictions is close to 2 percentage points—not a bad record. For winning parties the systematic error is +0.45. Thus the winner's percentages were *under*estimated by half a point on the average. Fortunately, the polls also underestimated the percentage of the second party by 0.29, which gave them a better record for their two-party predictions of −0.04—a trivial overestimate of the winner's percentage.

Table 1. Accuracy in Percentage Points of 155 National Poll Predictions, 1950–1984.[a] (Negatives are overestimates.)

	Winning Party	*Runner-up*	*Two-Party Split*
Average error	2.02	2.14	2.16
Systematic error	+0.45	+0.29	−0.04
Nonsystematic error	2.62	2.85	2.73

	Conservative	*Liberal/Labor*
Systematic error	+0.89	−0.08

	Systematic	*Nonsystematic*	*N*
By decade			
1950s	0.09	2.89	13
1960s	−0.39	2.15	29
1970s	0.15	2.60	67
1980s	−0.12	3.14	46

[a] Australia 20, Canada 16, France 26, West Germany 9, Japan 14, Netherlands 1, New Zealand 4, United Kingdom 36, United States 29.

The nonsystematic error is an indication of how closely these predictions cluster around the actual percentages, *after* allowing for these biases. This error component may be thought of as the net effect of all those uncorrected-for not-homes, nonvoters, undecideds, and mind-changers, plus random sampling errors, with some of the errors canceling out others. We may compare the standard deviations for these elections with the theoretical ones derived by the statisticians' formula, which represents the error due to sample size alone. Standard error from the formula above is $\sqrt{pq/n}$. The vote for the average winning party is 47 percent of the total, so we insert .47 for p and 1500 for n, our approximation of the average sample size, to get a theoretical standard error of 1.29 percentage points, which should be comparable to our nonsystematic errors.

But it is not. Standard deviation of the errors for winning parties turns out to be 2.62 percentage points, twice what would be expected. Multiplying by 1.96 we get a margin-of-error estimate of 5.1 percent, which we may compare to the *theoretical* margins of 2.5 percent to 3.5 percent as calculated above. Where theory would lead us to expect 7 or 8 errors larger than 3.5 percent, there were in fact 31 errors of this magnitude in the two-party predictions. As Paul Perry (1979) of the Gallup organization has observed: "One cannot expect to achieve in surveys an accuracy consistent with the theory of random sampling." On the basis of the long-term performance of the polling profession as a whole, press accounts of predictions should say "5 percent margin of error" where they now say "3 percent" and make comparable adjustments to all such estimates based on sample size.

Ideologues of left and right are convinced that pollsters rig their predictions to favor the other party. There is a greater likelihood, however, that the method itself is biased toward one side or the other. Perhaps the most perplexing problem faced in the adjustment decision is how to screen out those citizens who express an intention but are unlikely to reach the voting booth to implement it. Pollsters do this by eliminating or down-weighting on the basis of other questions about past voting habits, registration, interest in the election, and so on (Lipset, 1980; Traugott and Tucker, 1984; Wiener, 1976). If the screen is too loose, the intentions of nonvoters will be counted; if it is too tight some potential voters will be left out. There are also socioeconomic biases in sampling: e.g., the rich can restrict interviewers' access to their dwellings and can have unlisted phones; the poor are suspicious of interviewers and also may have no phones. The problem of election prediction is whether these screening, sampling, and other biases cancel out or whether they cumulate, as they obviously did in 1948.

We can shed some light on the question by grouping the parties into the more conservative ones (winners 103 times) and liberal or labor parties (winners 52 times) and comparing the systematic errors. It appears that there was an underestimate of conservative votes by nearly a percentge point (+0.89) and a trivial overestimate of the liberal/labor vote (−0.08). These led to predicting a liberal victory 17 times when the conservatives won (17 percent wrong) compared with 5 bad forecasts when the liberal or labor party won (10 percent wrong).

We might expect the predictions to have improved over the years since 1949, particularly in the 1980s, when random-digit-dialing and widespread telephone ownership have made possible forecasts based on data collected only a day or two before the election. Apparently not. The results by decades in Table 1 show no consistent pattern by either error measure.

In general the most serious errors occurred in a few elections in English-speaking democracies where several polls made the same mistakes. In the British election of 1970, five polls overestimated the Labor vote by an average of more than 4 percentage points, leading to a two-party error of some 3.5 points in favor of a party that lost the popular vote by 2.5 points. Thus 4 out of 5 forecasts picked the wrong winner. The year 1983 was in some respects even worse; the average of six polls was an overprediction of the Conservative vote by over 3 points and an *under*prediction of Labor by about 2 points, resulting in a two-party bias toward the Tories of nearly 4 points. But Mrs. Thatcher had such a commanding lead that no prediction was wrong and the error attracted less comment than the smaller one in 1970. Bad years in Australia were 1980, when a 3-point average underestimate of the Liberals caused five polls to go astray in a very close election, and 1984, when three out of four polls predicted a much wider margin for Labor than eventuated. In 1976 and 1980 in the U.S., both major parties were seriously underestimated, but the underestimates offset each other in the two-party percentages and put most of the polls on the right side. In all, 22 predictions (out of 155) pointed to the wrong winner, half of them occurring in these debacles.

The consistency with which almost all the polls in a few elections made errors in the same direction and of about the same magnitude is revealing. These are situations in which the electorate is volatile and people make up or change their minds in the last few days about whether and/or how they will vote. Survey procedures based on experience in more stable situations may not work as well. The popular stereotype that "the polls are always wrong" is incorrect. They are on the wrong side in only one prediction out of every seven. A better oversimplification would be: "Whenever the polls are wrong, they are all wrong."

To summarize, examining the performance of most of the published national election predictions since 1949 in a variety of political cultures and electoral systems shows: (1) The margin of error estimates based on probability theory seriously understate the empirical range of error. (2) Winners' percentages are underestimated on the average by about half a point and runners-up by a quarter point. (3) These compensating underestimates lead to better predictions of the two-party distribution and hence to picking the winner about 85 percent of the time. (4) There is no evidence that either biases or random errors have been reduced over the years. (5) There is a tendency to underestimate the conservative vote. (6) The bad predictions are made in a few elections when most of the polls err in the same direction.

References

Lipset, Seymour Martin
1980 "Different polls, different results in 1980 politics." Public Opinion, August/September, p. 19.

Mosteller, Frederick, Herbert Hyman, Phillip J. McCarthy, Eli S. Marks, and David B. Truman
1949 The Pre-election Polls of 1948. New York: Social Science Research Council.

Perry, Paul
1979 "Certain problems in election survey methodology." Public Opinion Quarterly 43:322–25.

Traugott, Michael W., and Clyde Tucker
1984 "Strategies for predicting whether a citizen will vote and estimation of political outcomes." Public Opinion Quarterly 48:330–43.

Weiner, Sanford L.
1976 "The competition for certainty: the polls and press in Britain." Political Science Quarterly 91:673–696.

Worcester, Robert M.
1983 Political Opinion Polling. New York: St. Martin's Press.

The Presence of Others and Overreporting of Voting in American National Elections

BRIAN D. SILVER, PAUL R. ABRAMSON
AND BARBARA A. ANDERSON

SURVEY RESPONDENTS often exaggerate their conformity with socially approved norms. To improve the quality of survey data, it is important to establish under what conditions people are more or less likely to exaggerate such conformity. Unfortunately, there are very few external standards against which reported behavior can be validated. Voting differs from most forms of political behavior because self-reports of voting can be compared with official voting statistics. Postelection sur-

---

Abstract This study uses the 1978 and 1980 vote validation studies conducted by the University of Michigan Survey Research Center to test the extent to which false claims about voting are affected by the presence of third parties during the interview. The presence of third parties during interviews is far more frequent than is commonly assumed. But the tendency of respondents to give socially approved answers is not affected by the presence of others during the interview. Thus, additional efforts to avoid contamination of interviews by eliminating third parties are not likely to reduce the exaggeration of self-reported vote. The analysis suggests that the declared intention to vote is a far more important factor in whether people falsely report voting than is the presence of others. Additional effort to understand the motivational basis of voting and nonvoting could help to account for variation in voting overreports.

Brian D. Silver and Paul R. Abramson are Professors of Political Science at Michigan State University, and Barbara A. Anderson is Professor of Sociology and Research Associate of the Population Studies Center at the University of Michigan. The data for this research were made available by the Inter-University Consortium for Political and Social Research, which bears no responsibility for the authors' analyses or interpretations. The authors are grateful to Mary R. Jackman, Robert W. Jackman, Howard Schuman, and Arland D. Thornton for helpful comments on an earlier draft; to Harriet Dhanak for programming assistance; and to Maria Elena Sanchez for providing information concerning the interviewer instructions in SRC National Election Studies. Work on this topic was motivated by research on the Soviet Interview Project, which is funded by Contract #701 from the National Council for Soviet and East European Research to the University of Illinois (James R. Millar, Project Director).

Public Opinion Quarterly Vol. 50:228–239 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-228/$2.50

veys in the United States consistently show a substantially higher proportion of the population claiming to have voted than the official vote records reveal.

In an attempt to determine the extent and patterns of this bias, the University of Michigan Survey Research Center conducted validation studies in 1964, 1976, 1978, and 1980, which allow comparison between self-reported voting and official voting records of the survey respondents. In these four studies, between 20 and 30 percent of the respondents who did not vote, according to the validation checks, claimed to have voted.

Although there have been several published articles devoted to the study of vote validation (e.g., Clausen, 1968–69; Katosh and Traugott, 1981; Sigelman, 1982), they have not shed much light on the conditions that contribute to false reporting of voting. Moreover, if respondents overreport their voting, they may overstate their participation in other socially desirable political activities, such as attending political meetings and contributing money to candidates. Although the actual extent of exaggeration by individuals of behaviors other than voting cannot be easily directly determined by comparisons with external criteria, understanding the sources of error in voting reports might lead to improved reliability of other reports as well.

Sudman and Bradburn (1974:1–27) and Bradburn (1983) identify three principal sources of variation in survey data quality: (1) the characteristics of the task (the survey instrument and context), (2) the interviewer's performance, and (3) the respondent's characteristics. Studies of vote validation have focused almost exclusively on the third source: respondent characteristics.[1] These studies have shown that overreporting of voting occurs in all social categories of the population, and they have concluded that the probability of overreporting is very weakly related to such respondent characteristics as socioeconomic status, age, and sex, or to such political attitudes as party identification, the sense of political efficacy, feelings of political trust, or sense of citizen duty.

The only social background characteristic that is reported to be strongly related to the size of the discrepancy between self-reported and validated vote is race: Blacks appear to be consistently more likely to overstate voting than whites (Traugott and Katosh, 1979; Katosh and Traugott, 1981; Sigelman, 1982; Abramson et al., 1983; Abramson and Claggett, 1984; Hill and Hurley, 1984).[2] However, given the far larger number of whites than of blacks, four-fifths of the respondents who overreport voting are whites.

---

[1] One exception is the research based on the Denver Validity Study of the late 1940s, which also examined interviewer effects. See Feldman et al., 1951; and Cahalan, 1968–69. Another is Weiss's (1968–69) analysis of the validity of voting reports (and other reports) of a sample of New York welfare mothers.

The main implication of previous research is that overreporting of voting is almost entirely a result of interviewer performance, characteristics of the survey task, or a combination of the two. Overreporting rarely results from a failure of memory, since the postelection surveys in which respondents are asked whether they voted in the recent presidential or midterm election are fielded within a few months of the election date, with a majority of the respondents interviewed within two months of the election. Instead, the response error appears to be more of the type obtained in market surveys in which people greatly exaggerate how frequently they brush their teeth (cf. Menneer, 1979:144).

There is some question, however, whether altering the characteristics of the survey task can reduce the tendency to exaggerate voting. There is very little published research on this topic. Research by Bradburn and Sudman (1980:1–13) suggests that altering the *form* of the survey, such as by using self-administered or telephone surveys or random response techniques, only slightly reduces the tendency of respondents to exaggerate their performance of socially desirable acts.

Other features of the survey task might contribute to the tendency toward overreporting of voting. If overreporting of voting is caused by the tendency of respondents to exaggerate their performance of socially desirable behavior, then the tendency toward exaggeration might be affected by whether people are present during the interview. The effect of the presence of third parties during interviews has rarely been studied.

Bradburn and Sudman report one study that compared the patterns of responses to attitudinal items between people who were interviewed in isolation and people who were interviewed in the presence of third parties. They found no consistent effect of third persons on the response patterns (Bradburn and Sudman, 1980:134–46). But Bradburn and Sudman's study did not test for the effects of third parties on responses whose accuracy could be *externally* validated. The SRC vote validation studies provide the opportunity for such a test.

---

[2] The Survey Research Center's validation of actual voting is not necessarily error-free. It is plausible that at least part of the discrepancy between reported and validated vote among blacks is due to errors in the records in local voting precincts, i.e., to the omission of some black voters. This seems especially likely to occur in rural areas and in the South. Southern *whites* also show a somewhat greater discrepancy between the number of self-reported and validated voters than do whites from other regions of the nation. See Abramson and Claggett, 1984.

How Is the Presence of Third Parties Likely to Affect Vote Overreporting?

Bradburn and Sudman (1980:134) state, "It is generally considered good field practice not to have anyone other than the interviewer and the respondent present during a face-to-face interview. . . ." One reason for trying to avoid having third parties present during interviews is to minimize contamination or distortion of the respondent's answers.

In the case of reported voting, the likely direction of bias caused by the presence of third parties is uncertain, because it may vary according to whether or not the third party knows whether or not the respondent did in fact vote. If the third party does not know whether or not the repondent voted, the respondent may be more likely to report socially desirable behavior and to claim to have voted even if he or she did not go to the polls. But in many cases the third party may know that the respondent did not vote, making it less likely for a nonvoter to claim falsely to have voted. In addition, parents may be especially reluctant to lie in front of their children.

The extent and direction of the effects of third-party presence is an empirical question, the answer to which rests on two behavioral conditions: (1) the frequency with which third parties are present during interviews, and (2) the impact of the presence of third parties on responses.

How Frequently Are Third Parties Present?

There is little published information about how often the desired goal of isolating the respondent one-on-one with the interviewer is met in practice. Bradburn and Sudman (1980) describe one national survey conducted in 1975 by the National Opinion Research Center (NORC) in which experienced interviewers were given special instructions to try to get all third parties out of the room before interviewing. In that survey, interviews were conducted with third parties present in 26 percent of the cases.

This suggests that when no special effort is made to reduce the presence of third parties, they are even more likely to be present. In Verba and Nie's (1972) *Participation in America* survey, conducted by NORC, third parties age 15 or older were present in 43 percent of the interviews.[3] In Jackman and Jackman's (1983) *Class Awareness in America* study, conducted by the SRC, third parties age 10 or older were present in 42 percent of the interviews.[4] In the Soviet Interview Project General Survey, conducted by NORC, interviews were conducted with third parties present in 52 percent of the cases (Anderson and Silver, 1986).

---

[3] This figure is calculated from the ICPSR codebook for the Verba-Nie study.

In surveys of women in less developed countries conducted for the World Fertility Survey, strict instructions were given to interviewers to administer the interviews privately (Casterline and Chidambaram, 1984:268). Nonetheless, on average in about half of the interviews third parties are reported to have been present.[5]

Unfortunately, few American survey research organizations routinely report information about the presence of third parties during the interviews. For example, NORC's General Social Survey does not gather this information. Fortunately, since 1966 the SRC American National Election Studies have collected information about the presence of others during interviews. This allows us to examine the relationship between overreporting of voting and the presence of third parties for both the 1978 and 1980 SRC Election Studies.[6]

Examination of the marginal distributions for those surveys reveals a very high incidence of third-party presence. As shown in Table 1, in the 11 SRC National Election surveys in which data on who was present during the interviews were collected, the lowest incidence of third parties being present was 40.6 percent (the 1980 postelection survey) and the highest was 56.8 percent (the 1970 postelection survey).

Despite the goal of isolating respondents one-on-one with the interviewers, third parties were present in the SRC National Election surveys during about *half* (51.2 percent) of the interviews. Consequently, if the presence of third parties during interviews affects the tendency to overreport socially desirable behavior, then the *potential* for an effect is considerably higher than is generally assumed.[7]

The effect of the presence of others might depend on the relationship

---

[4] Personal communication from Mary R. Jackman.

[5] On average, in World Fertility survey studies conducted in 24 less developed countries, third parties were present in 47.8 percent of the interviews (calculated from Table 1 of Casterline and Chidambaram, 1984).

[6] The 1964 Election Study conducted a vote validation but did not collect information about the presence of others in the interview. The 1976 Election Study collected information about the presence of others in the preelection survey but not in the postelection survey, when respondents were asked to report whether they had voted. Our treatment of the vote validation data from the 1978 and 1980 Election Studies follows the procedures used by Abramson and Claggett (1984).

[7] The bias would likely depend on whether third parties were present during the portions of the interviews when questions about socially desirable behavior were asked. The SRC National Election studies do not record when the third parties were present during the interviews. The instructions to interviewers state that third parties are to be counted as present if they were present during any part of the interview session.

Table 1. Presence of Others During the Interviews in SRC American National Election Studies, 1966–1982[a]

Who Was Present Other Than Respondent and Interviewer?	*1966 Post*	*1968 Pre*	*1968 Post*	*1970 Post*	*1972 Pre*	*1974 Post*	*1976 Pre*	*1978 Post*	*1980 Pre*	*1980 Post*	*1982 Post*
Nobody	44.3%	45.1%	49.3%	43.2%	44.8%	48.8%	54.3%	48.8%	49.7%	59.4%	55.3%
Children under 6 only	8.4	6.3	7.9	8.0	7.2	8.0	7.1	7.9	5.6	7.2	5.4
Older children or older & younger only	6.0	7.4	7.0	7.9	6.2	6.6	5.7	6.2	5.8	5.2	5.2
Spouse[b]	22.2	21.8	18.4	22.1	20.1	16.3	15.7	18.3	18.1	13.6	16.4
Spouse with children[c]	9.9	9.8	9.5	10.3	10.8	9.3	6.9	8.1	8.7	6.0	8.5
Other adults with or without children[d]	9.1	9.7	7.9	8.5	10.8	11.0	10.3	10.8	12.1	8.6	9.2
Total with somebody	55.7%	54.9%	50.7%	56.8%	55.2%	51.2%	45.7%	51.2%	50.3%	40.6%	44.7%
Base *N*[e]	(1279)	(1535)	(1334)	(1490)	(2690)	(2509)	(2847)	(2290)	(1613)	(1405)	(1416)

[a] Source: Compiled from frequency distributions in codebooks distributed by the Inter-University Consortium for Political and Social Research (Ann Arbor, MI). The table includes all SRC American National Election studies for which data on the presence of others were collected. The presidential year studies included both pre- and postelection interviews; the midterm election studies conducted postelection interviews only. However, information about the presence of third parties was not always collected in both the pre- and postelection surveys for a given year.

[b] Includes cases where other adults (relative or nonrelatives) were also present with spouse.

[c] Includes cases where other adults were also present with spouse.

[d] Includes cases where other relatives or other adults were present, with or without children, but without spouse being present.

[e] The Base *N* excludes respondents for whom the information was not ascertained. Figures for 1974 and 1976 are based on weighted *N*s.

of the third party to the respondent. As shown in Table 1, the most frequent "other person" in the interviews is the respondent's spouse; the second most frequent is a child. Table 2 shows that the probability that a third party will be present during the interview is about the same for blacks as for whites, and the same for males as for females. However, there are differences in the type of other person who is likely to be present.[8] The spouse is much more likely to be present for whites and for males than for blacks and for females. Children are more likely to be present for blacks and for females than for whites and for males.

Presence of Third Parties and Overreporting of Vote

Table 3 shows the percentage of validated nonvoters in 1978 and 1980 who say that they voted, controlling for the presence of other

---

[8] The SRC National Election surveys vary in size from year to year, partly because of the availability of funding and partly because of varying research goals. The 1978 midterm survey (which, like all but one midterm survey, was based only on a postelection interview) was the largest of the eight midterm surveys, mainly because of increased funding to study congressional elections. The 1980 survey (with a total *N* of 1614) was slightly smaller than the average for presidential election surveys, and our *N* is reduced further because the 206 respondents who were not included in the postelection interview are not included in our analysis.

Table 2. Presence of Others During Interviews, by Race and Sex, 1978 and 1980 SRC Studies[a]

Other Person Present?	*1978*				*1980*			
	Race		*Sex*		*Race*		*Sex*	
	White	*Black*	*Male*	*Female*	*White*	*Black*	*Male*	*Female*
Nobody	49.4%	46.8%	47.8%	49.8%	60.4%	53.0%	55.6%	62.3%
Child under 6	7.4	12.0	2.2	12.4	6.9	8.4	2.6	10.7
Older child	6.0	7.7	3.3	8.5	4.7	8.4	2.5	7.3
Spouse	19.4	8.6	25.0	12.8	14.3	8.4	20.9	7.9
Spouse & child	8.1	7.3	11.7	5.3	6.3	4.8	9.2	3.7
Other	9.7	17.6	10.0	11.2	7.3	16.9	9.3	8.1
Somebody (total)	50.6%	53.2%	52.2%	50.2%	39.6%	47.0%	44.4%	37.7%
Base *N*	(2003)	(233)	(1000)	(1263)	(1225)	(166)	(612)	(793)

[a] Source: Figures in this table and subsequent tables are calculated from data in the 1978 SRC American National Election Study and the 1980 American National Election Study postelection survey.

parties during the interview.[9] Nonvoters are slightly more likely to say that they voted when *nobody* other than the respondent and interviewer is present than when somebody else is present, but the difference is not statistically significant.[10] More important, the relationship between vote overreport and *who* was present shows little consistency between the 1978 and 1980 postelection surveys. For example, in the 1978 survey, validated nonvoters were *less* likely to say they voted when their spouse was present during the interview than when their spouse was not present; in the 1980 survey, validated nonvoters were *more* likely to say they voted when their spouse was present.

In both surveys, validated nonvoters were more likely to say they voted in the presence of older children than in the presence of younger children. It is possible that the presence of older children stimulated the respondents to provide a normatively correct response. Even if this is a systematic source of error, however, its aggregate impact is small, since, as Table 1 shows, older children are present in a very small proportion of the interviews. Moreover, the proportion of validated

---

[9] The *N*s for 1978 are substantially greater than those for 1980 for two reasons. First, as mentioned above, the 1978 sample is larger. Second, we are presenting the percentage who claimed to have voted only among validated *nonvoters*, and nonvoting is substantially higher in midterm elections than in presidential elections.

[10] A difference of proportions test (a two-tailed test) between whether Nobody was present and Somebody was present, based on the distributions reported in Table 3, revealed that the *p*-value for 1978 was .646 and the *p*-value for 1980 was .342.

Table 3. Percentage of Validated Nonvoters Who Say That They Voted, by Presence of Others During Interview, 1978 and 1980 SRC Postelection Surveys

		Other Persons Present							
Year		*Nobody*	*Somebody*	*Child Under 6*	*Older Child*	*Spouse*	*Spouse & Child*	*Other*	*All Nonvoters*
1978	%	23.1%	22.0%	22.9%	28.9%	19.5%	17.9%	23.9%	22.5%
	Base *N*[a]	(581)	(673)	(131)	(76)	(205)	(106)	(155)	(1254)
1980	%	29.2%	25.4%	15.5%	24.0%	34.4%	16.7%	31.5%	27.5%
	Base *N*[a]	(267)	(228)	(58)	(25)	(61)	(30)	(54)	(495)

[a] The Base *N* is the total number of validated nonvoters in each subset. The percentage shows the proportion of validated nonvoters who said that they voted.

nonvoters who report voting when children (young or old) are present is not consistently higher or lower than the average extent of overreporting by nonvoters.

There may be a consistent tendency to misreport less when *both* the spouse and child are present. This suggests that there may be a norm against lying in front of children. But the total number of interviews conducted with both the child and spouse present is small, and even the 11 percentage point difference in 1980 between this subset and all other validated nonvoters is not statistically significant. On balance, it appears safe to conclude that reports of voting by validated nonvoters are not affected very much by the presence of others during the interview.

Other Possible Factors Affecting Vote Overreporting

The lack of a consistent relationship between the kinds of people present and the extent of overstatement of voting is strong indirect testimony to the saliency of the civic norm of voting. Many Americans apparently find it difficult to acknowledge that they do not vote. There is at least indirect evidence in the SRC Election Study data that this reluctance to acknowledge not voting derives from the widespread norm that people should vote. One very strong predictor of whether voters were likely to report falsely that they voted—a predictor that has been overlooked in the research literature—is whether respondents stated in the preelection survey that they expected to vote.

Table 4 shows the percentage of validated nonvoters in 1980 who said that they voted, controlling for both the presence of others during the postelection interview and whether or not respondents said they expected to vote during the preelection interview. It shows a striking difference in the extent of vote overreporting between those who expected to vote and those who did not. This difference is far greater than

Table 4. Percentage of Validated Nonvoters Who Say That They Voted, by Whether Respondent Planned to Vote and by Presence of Others During Interview, 1980 SRC Post-Election Survey

		Other Persons Present							
Planned to Vote?[a]		*Nobody*	*Somebody*	*Child Under 6*	*Older Child*	*Spouse*	*Spouse & Child*	*Other*	*All Nonvoters*
Yes	%	53.7%	48.2%	31.8%	46.2%	59.4%	38.5%	53.1%	51.2%
	Base *N*[b]	(136)	(112)	(22)	(13)	(32)	(13)	(32)	(248)
No	%	1.0%	2.0%	0.0%	0.0%	7.4%	0.0%	0.0%	1.5%
	Base *N*[b]	(99)	(99)	(33)	(10)	(27)	(14)	(15)	(198)

[a] The question was worded as follows: "So far as you know how, do you expect to vote in November or not?" Cases where the response was coded Don't Know or Not Ascertained are excluded from the table.

[b] The Base *N* is the total number of validated nonvoters in each subset. The percentage shows the proportion of validated nonvoters who said that they voted.

the effects of the presence of third parties.[11] Respondents who said in the 1980 preelection interview that they planned to vote were much more likely to vote than those who said that they did not plan to vote. But among those who said they planned to vote and did not ($N = 248$), 51.2 percent said that they voted when interviewed after the election. Among respondents who said in the preelection survey that they did not plan to vote and who did not vote ($N = 198$), only three respondents said that they voted.

The consistency between preelection statements of *intention* to vote and postelection *reports* of voting supports the conclusion that variation in vote overstatement is caused in part by variation in the population in the commitment to the civic norm of voting.[12] Those who are committed to the norm of voting may feel that there is a stigma attached to reporting nonvoting to the interviewer, whereas those who are not committed to this norm are willing to report their nonvoting openly. The effect of these commitments is not substantially altered by the presence of third persons during the interview.

---

[11] There is some variation in overreporting related to who was present during the interview among nonvoters who said they expected to vote. But caution must be exercised not to overinterpret these apparent differences, especially given that the relationship between who was present and the reporting of voting by validated nonvoters differed sharply between 1978 and 1980.

[12] There is evidence that the preelection interview itself contributes to voting (Clausen, 1968–69); Traugott and Katosh, 1979), but this does not mean that the preelection statement of intention to vote contributes to the tendency of respondents in the postelection survey to claim falsely to have voted. As the 1978 vote validation study shows, there is substantial overreporting even when there is no preelection interview. Moreover, we have indirect evidence that the preelection statement that one plans to vote does not lead to an increase in the tendency to report that one voted after the election. In 1980, as in other presidential year surveys, most respondents were questioned by the same interviewer in both the pre- and postelection interview. If respondents falsely report voting in order to appear consistent, we would expect false reporting to be more frequent among those who were interviewed by the same interviewer in the pre- and postelection surveys. We compared the extent of false reports of voting among two subsets of validated nonvoters who said that they expected to vote: those who were questioned by the same interviewer and those who were questioned by different interviewers. We found no significant difference in reported turnout between these two subsets of validated nonvoters.

These results have more general implications for the analysis of overreporting of voting. While previous research has suggested that such overreports are not linked to social background characteristics of respondents (other than race) or to general political attitudes, overreporting may be partly a result of a more narrowly focused attitude toward the act of voting itself or toward the institution of elections (cf. Dennis, 1970).[13] If so, it is premature to conclude that respondent characteristics other than race do not account for vote overreporting.

Conclusion

We have shown that the presence of third parties during interviews is far more frequent than is commonly assumed. However, the presence of third parties does not explain why respondents say that they voted when they did not vote. There may be a tendency for nonvoters to be somewhat more honest in the presence of children, especially when their spouse is also present. This does not alter our overall conclusion that the tendency of many respondents to give the socially approved answer is not strongly affected by the presence of others during the interview.

The respondent's stated intention to vote, however, is a powerful predictor of subsequent voting overreports. Respondents who are positively inclined toward voting but who for some reason do not actually vote are likely to report falsely that they voted; respondents who are not inclined toward voting are not only unlikely to vote but are also highly unlikely to report falsely that they voted.

The results suggest that additional efforts to understand the motivational basis of voting and nonvoting could help to account for the extent of vote overreporting. But they also suggest that additional efforts to avoid contamination of interviews by eliminating third parties are not likely to reduce the exaggeration of self-reported vote.

The latter result is consistent with Bradburn and Sudman's (1980) conclusion that the presence of third parties has a minimal effect on survey response patterns. But it does not close the question whether it is desirable in general to have third parties present during interviews. There is a growing body of research on this topic that has not received sufficient attention from survey research specialists. In fertility surveys, studies of sex roles, and the study of the family, whether to interview spouses (or other family members) separately or together for the collection of both objective and subjective data is an important problem that demographers have addressed (e.g., Hermalin, 1975; Thomson, 1978; Thomson and Williams, 1982; Casterline and Chidambaram, 1984; Koenig et al., 1984).

---

[13] This interpretation is not consistent with the reported lack of an empirical relationship between the sense of citizen duty and vote overreporting. The citizen duty items all focus on people's sense of the importance of elections. But the statement of an intention to vote ("Do you expect to vote . . .?") may be a more valid reflection of a person's attitude toward the election than are the more general citizen duty items.

A general conclusion from the demographic literature is that there are sometimes advantages to having third parties present during interviews. This study cannot point to any advantages to having third parties present. But it does suggest that having third parties present during the interview does not lead to exaggeration of reported voting.

References

Abramson, Paul R., John H. Aldrich, and David W. Rohde

1983 Change and Continuity in the 1980 Elections, rev. ed. Washington, D.C.: Congressional Quarterly Press.

Abramson, Paul R., and William Claggett

1984 "Race-related differences in self-reported and validated turnout." Journal of Politics 46:719–38.

Anderson, Barbara A., and Brian D. Silver

1986 "The validity of survey responses: interviews of multiple respondents in a household from a survey of Soviet emigrants." University of Michigan, Population Studies Center, Research Report No. 86-89 (January).

Bradburn, Norman M.

1983 "Response effects," pages 289–328 in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research. New York: Academic Press.

Bradburn, Norman M., Seymour Sudman, and Associates

1980 Improving Interview Method and Questionnaire Design: Response Effects to Threatening Questions in Survey Research. San Francisco: Jossey-Bass.

Cahalan, Don

1968–69 "Correlates of respondent accuracy in the Denver validity survey." Public Opinion Quarterly 32:607–21.

Casterline, John B., and V. C. Chidambaram

1984 "The presence of others during the interview and the reporting of contraceptive knowledge and use," pages 267–98 in John A. Ross and Regina McNamara (eds.), Survey Analysis for the Guidance of Family Planning Programs. Liege, Belgium: ORDINA Editions.

Clausen, Aage R.

1968–9 "Response validity: vote report." Public Opinion Quarterly 32:588–606.

Dennis, Jack

1970 "Support for the institution of elections by the mass public." American Political Science Review 64:819–35.

Feldman, J. J., Herbert H. Hyman, and Clyde W. Hart
1951 "A field study of interviewer effects on the quality of survey data." Public Opinion Quarterly 15:734–61.
Hermalin, Albert I.
1975 "Models for the analysis of inconsistent reports of an event from two sources—with application to the accuracy of husband-wife reports." Presented to the Annual Meeting of the Population Association of America (Seattle, WA), April 17–19.
Hill, Kim Quaile, and Patricia A. Hurley
1984 "Nonvoters in voters' clothing: the impact of voting behavior misreporting on voting behavior research." Social Science Quarterly 65:199–206.
Jackman, Mary R., and Jackman, Robert W.
1983 Class Awareness in America. Berkeley and Los Angeles: University of California Press.
Katosh, John P., and Michael W. Traugott
1981 "The consequences of validated and self-reported voting measures." Public Opinion Quarterly 45:519–35.
Koenig, M. A., G. B. Simmons, and B. D. Misra
1984 "Husband-wife inconsistencies in contraceptive use responses." Population Studies 38:281–98.
Menneer, Peter
1979 "Retrospective data in survey research," pages 142–53 in Louis Moss and Harvey Goldstein (eds.), The Recall Method in Social Surveys. University of London Institute of Education, Studies in Education, new series 9.
Sigelman, Lee
1982 "The nonvoting voter in voting research." American Journal of Political Science 26:47–56.
Sudman, Seymour, and Norman M. Bradburn
1974 Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine.
Thomson, Elizabeth
1978 "The surrogate respondent: consistency of wives' and husbands' responses in a fertility survey and replication. Unpublished manuscript, Department of Sociology, University of Wisconsin.
Thomson, Elizabeth, and Richard Williams
1982 "Beyond wives' family sociology: a method for analyzing couple data." Journal of Marriage and the Family 44:999–1008.
Traugott, Michael W., and John P. Katosh
1979 "Response validity in surveys of voting behavior." Public Opinion Quarterly 43:359–77.
Verba, Sidney, and Norman H. Nie
1972 Participation in America: Political Democracy and Social Equality. New York: Harper and Row.
Weiss, Carol H.
1968–9 "Validity of welfare mothers' interview responses." Public Opinion Quarterly 32:622–33.

Opinions on Fictitious Issues: The Pressure to Answer Survey Questions

GEORGE F. BISHOP, ALFRED J. TUCHFARBER
AND ROBERT W. OLDENDICK

A VARIETY of studies since the early years of public opinion research have demonstrated that people are quite willing to express opinions on fictitious issues, objects, and events.[1] Bishop and his associates (1980,

---

[1] For an early example, see Hartley's (1946) studies of prejudice in which large majorities of college students were willing to express opinions about fictitious nationalities (e.g., "Wallonians"). See also Kolson and Green's (1970) demonstration of opinions among grade-school children about a nonexistent political figure (Thomas Walker) and Collett and O'Shea's (1976) study of people giving directions to a fictional place. References are frequently made, too, to Gill's (1947) report of large numbers of people who were willing to offer opinions on a fictional "Metallic Metals Act," but Schuman and Presser (1981) have pointed out that this report is just an anecdote with no documentation provided on the population or the sample size.

Abstract This research shows, as have previous studies, that a substantial number of people will offer opinions on fictitious topics in the context of a survey interview. The results of a split-ballot experiment (and a replication) with three fictitious public affairs issues suggest that people give opinions on such matters in large part because of the pressure to answer survey questions that is created by the way in which the questions are asked and by the manner in which "don't know" responses are handled by the interviewer. The findings also indicate that there may be little or no relationship between an individual's willingness to admit ignorance and his or her inclination to volunteer opinions on fictitious issues, as measured by at least one item from the Marlowe-Crowne Social Desirability Scale. In addition, the data demonstrate that blacks and less well-educated respondents are more likely than their counterparts to offer opinions on fictitious issues.

George F. Bishop is Associate Professor of Political Science and a Senior Research Associate in the Institute for Policy Research at the University of Cincinnati. Alfred J. Tuchfarber is Associate Professor of Political Science and Director of the Institute for Policy Research. Robert W. Oldendick is Assistant Director of the Institute for Policy Research. This research project was supported by a grant from the Procter & Gamble Company of Cincinnati, Ohio. The authors would like to thank Jack Henry, Manager of Marketing Research at P&G, for funding the project. The authors would also like to thank the anonymous referees for their helpful comments on an earlier version of the manuscript.

Public Opinion Quarterly Vol. 50:240–250 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-240/$2.50

1983), for example, have found that a sizable percentage of respondents (roughly 30 percent) will give an opinion on a fictitious "1975 Public Affairs Act" if they are not offered an explicit response alternative which allows them to admit that they "don't know" anything about it. Similarly, Schuman and Presser (1981) have discovered that about 30 percent of the respondents in their national surveys will, if not given an explicit chance to say thay have "no opinion" on the subject, answer questions about real, but highly obscure, pieces of legislation such as The Agricultural Trade Act of 1978 and The Monetary Control Bill of 1979. Such are the pressures in a typical interview to answer survey questions.

From such studies, we have also learned that some people are more likely than others to offer opinions on obscure or fictitious issues: blacks more likely than whites, and the less educated more likely than the better educated (Bishop et al., 1980; Schuman and Presser, 1981; Ch. 5). The latter finding is somewhat puzzling because, as Schuman and Presser have noted, less educated respondents are generally *less likely,* not *more likely,* to volunteer opinions on public affairs issues. What explains the difference? Schuman and Presser have suggested that better educated people may be more willing than those who are less well-educated to admit their ignorance on topics which they do not recognize. Bishop and his colleagues (1980) have proposed a similar explanation, hypothesizing that blacks and less educated respondents, of whatever race, may offer opinions on the Public Affairs Act to "save face" with "our presumptively white middle-class interviewing staff . . . calling from the University. . . ."[2] Common to both these explanations is the idea that people have a disposition or personality characteristic which makes them either more or less willing to admit their ignorance on public affairs.

Neither of these dispositional hypotheses, however, has been directly tested or pitted against an equally plausible explanation, also suggested by Schuman and Presser: "that low-educated respondents are especially susceptible to pressures in the interview to give an opinion even where they have none (p.152)." Furthermore, Schuman and Presser have noticed that, in some instances, interviewers increase the pressure on respondents to give an answer by (inappropriately) repeating the question to them even though they have already indicated that they "don't know" anything about the subject. Respondents thus get the message that the interviewer *expects* them to answer the question. The majority of respondents do not conform to this implicit demand, but a sizable number do, indicating that the pressure to give an opinion in this situation can be quite compelling. That, in fact, is what the experiment described here was designed to demonstrate: that people give opinions on obscure or fictitious issues largely because of the "demand characteristics" of the survey interview (cf. Orne, 1962). The experiment was also designed to test the alternative dispositional hypothesis: that people give opinions on such matters because they are unwilling to admit that they are ignorant. Included in the design, as well, was an analysis of the response effects of the two major demographic variables in previous research: race and education.

---

[2] Bishop et al. (1980, fn. 10) report an almost perfect correlation (.99) between the race of the respondent perceived by the interviewer in the first wave of a telephone panel study and the respondent's self-reported race in the second wave, indicating that the race of the respondent (and by implication, the interviewer's) can be accurately inferred over the phone.

Research Design

In November and early December of 1983 we conducted the first of two split-ballot experiments on people's responses to questions about three fictitious issues: The 1975 Public Affairs Act, The Agricultural Trade Act of 1983, and The Monetary Control Bill of 1983. The questions were all asked, together, near the middle of the interview, as part of a multipurpose, random-digit-dialed telephone survey,[3] and either immediately before or after questions about three public affairs issues which were not fictitious (see Appendix). The latter items were included to provide a cover for asking respondents about the fictitious issues as well as a test and control for the influence of question context.[4]

The question about the Public Affairs Act was drawn from the experiments by Bishop et al. (1980). The questions about the Agricultural Trade Act and the Monetary Control Bill were adapted from the items developed originally by Schuman and Presser (1981; Ch. 5): The Agricultural Trade Act of *1978* and The Monetary Control Bill of *1979* (emphais added). We changed the year of these legislative acts—to

---

[3] The response rate (fully completed interviews) for this Greater Cincinnati Survey was 71.5 percent. The refusal rate was 15.8 percent. The remainder consisted of partially completed interviews (3.2 percent) and potential interviews that were not completed because of a language barrier, a hearing problem, illiteracy, senility, or physical illness (5.6 percent) or because the selected respondent was away on vacation, a business trip, or avoided an appointment for some other reason (3.9 percent).

[4] The question about whether the federal government is doing too many things that should be handled by the private sector was adapted from an item developed originally by the National Opinion Research Center (see NORC Study 4179, December 1973; 14, University of Chicago). The questions on trade with the Soviet Union and government spending were both adapted from items designed by the Center for Political Studies at the University of Michigan (see the 1982 National Election Study Cross-Section Interview: 52 and 55).

1983—to make them appear up-to-date and thus more compelling as issues on which the respondent should probably have an opinion.[5] To our knowledge, there were no such legislative matters pending in Congress at the time of survey. Nor is it likely that any respondent knew anything about the real, but obscure, legislation from which these items were initially derived.

QUESTION FORMS

The questions were asked in one of three forms, to which respondents were randomly assigned:

1. *Form A*. In this form the questions were asked with an explicit response alternative (a filter) which allowed respondents to indicate that they hadn't thought much about the issue (see Appendix).[6]

2. *Form B*. Here the questions were asked without an explicit filter. But if the respondents volunteered that they didn't know anything about the topic or gave a qualified answer to the question other than the two response alternatives that were read to them (e.g., agree/disagree), the interviewers were instructed *not to probe* them in any way, but rather to record, verbatim, whatever the respondent said.

3. *Form C*. The questions in this form were also asked without a filter. However, if the respondents volunteered a "don't know" response or offered a qualified answer to the question, the interviewers were told to pressure them into selecting one of the two forced-choices (e.g., favor/oppose) by repeating the question, by rereading the choices, or both.

Forms B and C were varied to test Schuman and Presser's proposition about the pressure interviewers put on respondents by probing, or what we call the "demand characteristics" hypothesis.

QUESTION ORDER AND CONTEXT

In addition, we varied the order in which the groups of questions were asked, randomly assigning respondents to one of two conditions: in one condition, respondents were asked the group of questions about the fictitious issues immediately *before* the group of questions about the nonfictitious issues; in the other condition, immediately *after* them. Within each of these groups, however, the three questions were asked in the same sequence in both conditions (see Appendix). This manipulation allowed us to test whether respondents would be more susceptible to being pressured into answering the questions about the fictitious issues when they were asked in one context rather than another.

---

[5] The wording of the filter used with our adaptation of these questions—i.e., "or haven't you thought much about this issue?"—differs from the one used by Schuman and Presser—i.e., "or do you not have an opinion on that issue?" This variation may have had some effect on the proportion of DK responses observed in our experiment, but it is probably a minor one (see Bishop et al., 1983 on the effects of differently worded filter questions).

[6] Respondents receiving this form were *not* probed if they volunteered a "don't know" response or gave a qualified answer to the item. This, we reasoned, was consistent with the intent of the filtered form, which was to make it acceptable not to have an opinion on an issue. It also seemed consistent not to pressure respondents who received this form into selecting one of the forced-choices, but rather to let them "freely" qualify their answers if they wanted to.

MEASUREMENT OF WILLINGNESS TO ADMIT IGNORANCE

To tap the disposition to admit ignorance, we adapted an item from the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1964) and asked it along with four other items from the same scale which allowed us to construct a general index of the tendency of people to present themselves to others in a favorable light. The five items, the fourth of which measured a person's willingness to admit ignorance, were worded as follows.[7]

> Now I'd like to read you some statements about personality characteristics. As I read each one, please tell me whether the statement is true or false as it applies to you personally.
>
> 1. I have never deliberately said something that hurt someone's feelings—is that true or false?
> 2. I'm always willing to admit it when I make a mistake—is that true or false?
> 3. No matter who I am talking to, I'm always a good listener—is that true or false?
> 4. When I don't know something I don't at all mind admitting it—is that true or false?
> 5. I'm always courteous, even to people who are disagreeable—is that true or false?

Because responses to the question about admitting ignorance might influence responses to questions about the fictitious issues, or vice versa, we decided to randomize the location of the Marlowe-Crowne items: they were asked shortly before the fictitious and nonfictitious issues in one condition; shortly after them, in the other. To hide the connection between the question about admitting ignorance and the fictitious issues, we not only embedded it among the other Marlowe-Crowne items but also created a brief "buffer" between these items and the issues with two questions about the local major league baseball team. The latter were asked immediately *after* the Marlowe-Crowne items, and just *before* the issue questions, in one condition; immediately *before*, and just *after*, in the other.

---

[7] If respondents tried to qualify their answers to these questions by saying such things as, "Well that's true sometimes, but not always," interviewers were instructed to probe them by asking: "Is it *mostly true* or *mostly false*?" In addition, they were instructed to probe all "don't know" responses by repeating the question. This was done to minimize missing data.

REPLICATION

In May and early June of 1984 we replicated the entire experiment, exactly, again as part of a multipurpose, random-digit-dialed telephone survey.[8]

Findings

The data in Table 1 demonstrate that the form in which the questions are asked, and the way in which the "don't know" (DK) responses are dealt with, makes a substantial difference in the results.[9] People were much more likely to offer an opinion on a fictitious issue when the question was asked without an explicit filter (Forms B and C) than when it was asked with one (Form A). And they were even more likely to offer an opinion on the unfiltered form when "don't know" responses were probed (Form C) than when they were not (Form B). This pattern occurred on all three of the issues, and in both surveys.[10]

Notice, however, that respondents were generally more likely to offer opinions on both the Agricultural Trade Act and the Monetary Control Bill than they were on the Public Affairs Act, probably because they were more likely to confuse the former topics with subjects with which they were familiar, such as the questions we asked about trade with the Soviet Union and government spending (see Appendix). People can thus be more easily manipulated into giving opinions on subjects which seem at least faintly familiar than those which do not.

EFFECTS OF WILLINGNESS TO ADMIT IGNORANCE

We found little or no relationship between an individual's willingness to admit ignorance and his or her tendency to offer opinions on

---

[8] The response rate for this Greater Cincinnati Survey was 71.3 percent. The refusal rate was 14.8 percent. The rest consisted of partially completed interviews (2.8 percent) and potential interviews that were not completed because of a language barrier, a hearing problem, illiteracy, senility, or physical illness (5.6 percent) or because the selected respondent was away on vacation, a business trip, or avoided an appointment for some other reason (5.5 percent).

[9] Because there were no significant differences in responses to the fictitious issues, in either of the two surveys, when the questions were asked before or after the nonfictitious issues (data not shown here), we have combined the data for these two context conditions in constructing Table 1.

[10] The differences between Forms B and C tend to be somewhat greater in the May/June 1984 survey, especially on the Public Affairs Act, indicating that the interviewers may have pressured respondents more heavily and consistently, perhaps because of greater emphasis in training for this survey on how to probe respondents who give "don't know" responses.

Table 1. Percent with Opinion on Public Affairs Act, Agricultural Trade Act, and Monetary Control Bill by Question Form and Survey

	Filtered	*Unfiltered*	
	Form A (DK's Not Probed)	*Form B (DK's Not Probed)*	*Form C (DK's Probed)*
Public Affairs Act	3.7	24.3	31.5
(Nov./Dec., 1983)[a]	(408)	(395)	(399)
Public Affairs Act	3.6	22.2	42.3
(May/June, 1984)[b]	(397)	(411)	(402)
Agricultural Trade Act	12.1	38.9	49.0
(Nov./Dec., 1983)[c]	(405)	(399)	(398)
Agricultural Trade Act	10.5	35.5	55.9
(May/June, 1984)[d]	(398)	(411)	(403)
Monetary Control Bill	13.2	33.7	46.3
(Nov./Dec., 1983)[e]	(408)	(397)	(397)
Monetary Control Bill	14.3	38.4	54.7
(May/June, 1984)[f]	(398)	(411)	(404)

[a] Form A vs. B vs. C: $\chi^2=106.05$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=4.86$, df=1, $p<.03$).

[b] Form A vs. B vs. C: $\chi^2=171.01$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=36.93$, df=1, $p<.0001$).

[c] Form A vs. B vs. C: $\chi^2=130.21$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=7.91$, df=1, $p<.01$).

[d] Form A vs. B vs. C: $\chi^2=183.33$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=33.02$, df=1, $p<.0001$).

[e] Form A vs. B vs. C: $\chi^2=104.05$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=12.57$, df=1, $p<.001$).

[f] Form A vs. B vs. C: $\chi^2=143.68$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=21.00$, df=1, $p<.0001$).

fictitious issues, regardless of the form in which the questions were asked (data not shown here).[11] The failure to find such a relationship does not, however, mean that the dispositional hypothesis can be discarded because the measure we used was admittedly a rather weak one, based on a single item, the responses to which were highly skewed: over 90 percent of the respondents in the first survey, and nearly 90 in the second, said they don't at all mind admitting it when they "don't know" something. So there was little or no variance in the item, and therefore little opportunity for it to differentiate between respondents who give opinions on fictitious issues and those who do not. Perhaps, then, a better measure, or set of measures, of a person's willingness to admit ignorance would result in a different conclusion.

---

[11] Nor did we find any significant relationship, in either study, between a respondent's score on an overall index of social desirability—formed by summing responses to all five of the Marlowe-Crowne items—and his or her response to any of the fictitious issues (data not shown).

Table 2. Percent with Opinion on Public Affairs Act, Agricultural Trade Act, and the Monetary Control Bill by Race, by Education, by Question Form (Fall 1983 and Spring 1984 Surveys Combined)

	Filtered Form (A)				*Unfiltered Forms (B & C Combined)*			
	Better Educated		*Less Well Educated*		*Better Educated*		*Less Well Educated*	
	Blacks	*Whites*	*Blacks*	*Whites*	*Blacks*	*Whites*	*Blacks*	*Whites*
Public Affairs Act[a]	4.0	5.0	4.2	2.2	41.5	22.1	52.2	30.9
	(50)	(299)	(71)	(368)	(82)	(633)	(159)	(706)
Agricultural Trade Act[b]	16.0	11.4	4.3	12.2	51.8	38.4	59.0	46.5
	(50)	(297)	(70)	(369)	(83)	(635)	(156)	(710)
Monetary Control Bill[c]	16.0	18.5	16.9	9.5	56.6	40.8	59.5	40.5
	(50)	(298)	(71)	(370)	(83)	(633)	(158)	(709)

NOTE: Better educated = some college or more; less well-educated = high school education or less.

[a] Differences on filtered form *not* statistically significant. Differences on unfiltered forms: better educated blacks vs. whites ($\chi^2 = 13.72$, df = 1, $p < .001$); less well-educated blacks vs. whites ($\chi^2 = 25.07$, df = 1, $p < .0001$).

[b] Differences on filtered form *not* statistically significant. Differences on unfiltered forms: better educated blacks vs. whites ($\chi^2 = 4.93$, df = 1, $p < .03$); less well-educated blacks vs. whites ($\chi^2 = 7.50$, df = 1, $p < .01$).

[c] Differences on filtered form *not* statistically significant. Differences on unfiltered forms: better educated blacks vs. whites ($\chi^2 = 6.92$, df = 1, $p < .01$); less well-educated blacks vs. whites ($\chi^2 = 18.20$, df = 1, $p < .0001$).

EFFECTS OF RACE AND EDUCATION

As expected from previous research, both race and education had a significant influence on responses to the fictitious issues, depending upon the form in which the questions were asked (see Table 2).[12] On the filtered form (A), the differences between blacks and whites, and between the better educated and the less well educated, were not statistically significant on any of the three times. Most respondents, in fact, regardless of their race or education, said they hadn't thought much about any of the fictitious issues when asked about them. On the unfiltered form, however, the differences between groups were sizable as well as statistically significant. Blacks offered opinions on all three issues significantly more often than whites, whether they were better educated or not. Notice, too, that education had an independent effect, particularly on responses to the Public Affairs Act: the less well educated offered opinions on the issues more often than the better educated, whether they were black or white. But the effects of education were not as strong as those of race; in fact, on the Monetary Control Bill, there was no significant difference in response by education when race was held constant. Race is thus the more important of the two.[13]

---

[12] Because of small cell sizes in the black subsample, the data for the two surveys have been combined and education has been collapsed into two categories, an admittedly crude control, but inescapable under the circumstances.

Conclusion

The findings suggest that people give opinions on fictitious issues, in large part, because of the *pressure to answer* which is created by the form in which the questions are asked and the manner in which the "don't know" responses are handled by the interviewer. The data also indicate that people can be more easily pressured into giving an opinion on a fictitious issue when the topic seems familiar to them (e.g., agricultural trade), probably because they confuse it with a related subject on which they do have an opinion. The more a person knows about a subject, however (e.g., politics, sports), the less likely he or she is to make such a mistake; in fact, the more they know about the subject, the more easily they can recognize what is familiar and what is not (e.g., a fictional politician or baseball team). Blacks and less well-educated respondents may thus be more likely to offer opinions on fictitious issues because they are generally less well informed about politics and therefore less able to tell whether the topic is something that "belongs."[14] As a general principle, then, we would hypothesize that the less knowledgeable a person is about a given subject, the more easily he or she can be confused and pressured to give an opinion about it. Such are the "demand characteristics" of the survey interview.

What are the implications of these findings for the practice of probing? Surely we should not pressure people to give opinions when they plead ignorance. Most of the problem can, in fact, be avoided by giving respondents an explicit opportunity to say they have "no opinion" on the subject—for example, by using a filter question. But on many public affairs issues this will mean losing 20 to 25 percent of the cases for the analysis (see Bishop et al., 1983; Schuman and Presser, 1981:

---

[13] We also discovered that, with the exception of the Public Affairs Act, men were somewhat more likely than women to offer opinions on the fictitious issues on both the filtered and unfiltered forms, but the differences were relatively small—less than 10 percent on the average (data not shown). Younger people were significantly more likely than older people to give opinions on the Monetary Control Bill when the question was asked in unfiltered form, but other than that there were no significant differences among age groups.

[14] We have found in a previous survey, for example, that blacks are significantly less likely than whites to know the names of their United States senators, the length of a senator's term of office, and which political party has the most members in the U.S. House of Representatives (data not presented here).

Ch. 4). That, for many pollsters, however, may be too big a price to pay in the loss of representativeness. Nor does it make good copy to say that many people have no opinion on the issues of the day. The alternative is to create a spurious form of representativeness by minimizing the "don't knows." Perhaps we should stop probing altogether, except to ask respondents what they mean when they say "I don't know." For probing presumes that the respondent has an opinion on the issue that merely needs to be drawn out. Yet, how can we ever know that in advance? Is it not time, then, that we reexamined the practice of probing in public opinion surveys? And with it, the meaning of what we so routinely record as "DK."

Appendix

WORDING OF THE QUESTIONS ON FICTITIOUS AND NONFICTITIOUS ISSUES

Form A	*Forms B and C*
Some people feel that The 1975 Public Affairs Act should be repealed—do you agree or disagree with this idea, or haven't you thought much about this issue?	Some people feel that *The 1975 Public Affairs Act* should be repealed—do you agree or disagree with this idea?
Congress has been considering *The Agricultural Trade Act of 1984*—do you favor or oppose the passage of this act, or haven't you thought much about this issue?	Congress has been considering *The Agricultural Trade Act of 1984*—do you favor or oppose the passage of this act?
Congress has also been considering *The Monetary Control Bill of 1984*—do you favor or oppose the passage of this bill, or haven't you thought much about this issue?	Congress has also been considering *The Monetary Control Bill of 1984*—do you favor or oppose the passage of this bill?
. . . Do you think the government is trying to do too many things that should be left to individuals and private businesses, do you think the government should do even more to solve our country's problems, or haven't you thought much about this issue?	. . . Do you think the government is trying to do too many things that should be left to individuals and private businesses, or do you think the government should do even more to solve our country's problems?
. . . Do you think the United States should do more trading and business with the Soviet Union, do you feel that the United States should cut back on its economic ties with the Soviet Union, or haven't you thought much about this issue?	. . . Do you think the United States should do more trading and business with the Soviet Union, or do you feel that the United States should cut back on its economic ties with the Soviet Union?

Form A (cont.)	*Forms B and C* (cont.)
. . . Do you think the federal government should provide fewer services, even in areas such as health and education, in order to reduce spending, do you feel it is important for the government to provide many more services even if it means an increase in spending, or haven't you thought much about this issue?	. . . Do you think the federal government should provide fewer services, even in areas such as health and education, in order to reduce spending, or do you feel it is important for the government to provide many more services even if it means an increase in spending?

References

Bishop, George F., Robert W. Oldendick, and Alfred J. Tuchfarber
1983 "Effects of filter questions in public opinion surveys." Public Opinion Quarterly 47:528–46.
Bishop, George F., Robert W. Oldendick, Alfred J. Tuchfarber, and Stephen E. Bennett
1980 "Pseudo-opinions on public affairs." Public Opinion Quarterly 44:198–209.
Collett, Peter, and Gregory O'Shea
1976 "Pointing the way to a fictional place: a study of direction giving in Iran and England. European Journal of Social Psychology 6:447–58.
Crowne, Douglas P., and David Marlowe
1964 The Approval Motive. New York: Wiley.
Gill, Sam
1947 "How do you stand on sin?" Tide magazine, March 14, pp.72 ff.
Hartley, Eugene L.
1946 Problems in Prejudice. New York: Octagon Press.
Kolson, Kenneth L., and Justin J. Green
1970 "Response set bias and political socialization research." Social Science Quarterly 51:527–38.
Orne, Martin J.
1962 "On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications." American Psychologist 17:776–83.
Schuman, Howard, and Stanley Presser
1981 Questions and Answers in Attitude Surveys. New York: Academic Press.

Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys

ROBERT M. GROVES AND LOU J. MAGILAVY

INTERVIEWER variance or interviewer effects reflect the tendency for answers provided by the respondent and recorded in a questionnaire to vary depending on which interviewer is assigned to the respondent. In order to obtain estimates of these effects, "interpenetrated" designs are typically introduced, which randomly assign sample cases to interviewers. Although several researchers have touted the possibility of cheap interpenetration in centralized telephone surveys (see Groves and Kahn, 1979; Tucker, 1983), routine measurement has been found to be a greater challenge than previously assumed. This article sum-

---

Abstract Estimates of interviewer effects on survey statistics are examined from nine surveys conducted over a six-year period at the Survey Research Center. Estimates of intraclass correlations associated with interviewers are found to be unstable, given the number of interviewers (30–40) used on most surveys. This finding calls into question inference from earlier studies of interviewer effects. To obtain more reliable information about magnitudes of interviewer effects, generalized effects are constructed by cumulating estimates over statistics and surveys. These generalized correlations are found to be somewhat smaller than those reported in the past literature. Few differences in generalized interviewer effect measures are found between open and closed questions or between factual and attitudinal questions. Small reductions in effects *were* obtained when a Computer Assisted Telephone Interviewing (CATI) system was used; there was some evidence of elderly respondents being more susceptible to interviewer effects; the number and type of second responses to open questions were affected by interviewer behavior; and changes in interviewing techniques reduced interviewer effects.

Robert M. Groves is an Associate Research Scientist in the Survey Research Center and an Associate Professor in the Department of Sociology at The University of Michigan. Lou J. Magilavy is a Ph.D. candidate in the Department of Biostatistics at The University of Michigan. Data in this article were collected in projects supported by the National Science Foundation and the National Center for Health Statistics. The views in this article do not necessarily represent those of either organization. An earlier version of this article was presented at the meetings of the International Statistical Institute, 1983. The authors appreciate the comments of Leslie Kish, Stanley Presser, and Howard Schuman on earlier versions, and extend their gratitude to Richard Curtin for access to the data from the Surveys of Consumer Attitudes.

Public Opinion Quarterly Vol. 50:251–266 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-251/$2.50

marizes six years of research in the measurement of interviewer effects in centralized telephone surveys, on topics ranging from health experiences to political attitudes. We attempt to address two weaknesses in past research—problems of estimation (obtaining precise and accurate estimates of interviewer effects) and problems of explanation (learning what design features affect the magnitude of interviewer effects).

Study Designs

Eight of the nine surveys examined used a two-stage stratified national sample of randomly generated telephone numbers, following the design of Waksberg (1978). The ninth survey, the 1980 Post Election Study, was a telephone reinterview with respondents chosen earlier in an area probability sample. In all surveys one adult was randomly selected in each sample household. As shown in Table 1, response rates ranged from 59 percent to 87 percent, with unanswered numbers included in the denominators of these rates.[1] Row 4 lists the number of interviews obtained from the random assignment of telephone numbers to interviewers. Interviews resulting from nonrandom assignments (e.g., refusal conversions, select appointments) were excluded from both those numbers and the following analyses. Also presented is the number of interviewers employed for each study and the average workload (i.e., number of interviews) per interviewer. Not only does average workload vary across studies, from 12 for the November 1981 Survey of Consumer Attitudes to 58 for the Health in America Survey, but workloads within studies exhibit a similar degree of variation. All the studies collected standard demographic data in addition to data on the specific topics listed in Table 1. The 30 variables analyzed for the Survey of Consumer Attitudes are identical in each of the five months. The number of statistics analyzed for each study ranges from 25 to 55.

Estimation of Interviewer Effects

THE INTERVIEWER EFFECTS ESTIMATOR

To compare interviewer effects across statistics and to maximize comparability to past interviewer variance research we utilize an intraclass correlation, based on a one-way analysis of variance, using interviewers as the factor, $\rho_{int}^{*} = S_a^2/(S_a^2 + S_b^2)$, where S_a^2 is the between interviewer variance and S_b^2 is the within interviewer variance, with

---

[1] The lowest response rate, 59 percent, reported for the Study of Telephone Methodology reflects an exceptionally large proportion of unanswered numbers. In later studies, improved methods of identifying nonworking numbers have caused a decrease in this proportion and consequently a more accurate response rate.

Table 1. Study Descriptions and Summary of Overall Findings

Study Description and Summary Statistics	*Study of Telephone Methodology*	*Health and Television Viewing*	*Health in America*	*1980 Post Election Study*	*Monthly Survey of Consumer Attitudes*				
					Nov '81	*Dec '81*	*Jan '82*	*Feb '82*	*Mar '82*
Field period	Apr– May 1976	Mar– Apr 1979	Oct– Dec 1979	Nov– Dec 1980	Nov '81	Dec '81	Jan '82	Feb '82	Mar '82
Response rate	59%	67%	80%	87.3%	73.5%	72.8%	73.3%	76.9%	73.8%
No. interviewers	37	30	33	22	31	28	21	25	26
No. interviews	1529	954	1918	697	370	350	386	379	366
Average workload	41.3	31.8	58.1	31.7	11.9	12.5	18.4	15.2	14.1
Major topics	Political economic & social issues	Health and TV viewing	Health	Political issues	Economic issues	Economic issues	Economic issues	Economic issues	Economic issues
No. variables analyzed	25	55	25	42	30	30	30	30	30
Range of ρ^*_{int}	$-.0080$	$-.0150$	$-.0070$	$-.0154$	$-.0217$	$-.0373$	$-.0221$	$-.0419$	$-.0356$
	.0560	.1650	.0097	.1710	.0895	.0546	.0916	.0657	.0729
Mean values									
ρ^*_{int}	.0089	.0074	.0018	.0086	.0184	.0057	.0163	.0090	.0067
$deff_{int}$	1.36	1.23	1.10	1.26	1.20	1.07	1.28	1.13	1.09

large positive values of ρ^*_{int} reflecting large interviewer effects. This is the measure described more fully by Kish (1962) and used frequently in this literature. His studies of factory workers' job attitudes produced ρ^*_{int}'s of $-.03$ to .09 (mean = .02). O'Muircheartaigh and Marckwardt (1980) presented fertility statistics most susceptible to interviewers, with ρ^*_{int}'s between .00 and .19 (mean = .07). Feather (1973) in a health survey found a range of ρ^*_{int}'s of $-.01$ to .03 (mean = .006). Freeman and Butler (1976), in a study of attitudes toward mental retardation, displayed unusually high values of ρ^*_{int} (mean = .04), that might be attributable to the use of nonprofessional interviewers. The overall average for all these statistics from personal interview surveys is .018. Tucker (1983) presented mean ρ^*_{int}'s from 11 telephone surveys ranging from $-.003$ to .008 (overall mean = .004) for political and social attitudes.[2]

For this research we estimated ρ^*_{int} after interpenetrating the sample within shifts of a centralized telephone interviewing facility. This randomization of assignment within shift is an added constraint on the model, the effect of which was judged to be negligible in these studies.[3]

It is important to realize that the effect of the interviewer on the overall variance of survey statistics is a function of ρ^*_{int} *and* the size of the interviewer workload, *m*. The variance of a survey statistic is inflated by the factor $deff_{int} = [1 + \rho^*_{int}(m - 1)]$. Thus, if interviewers complete large numbers of cases, say 50, even ρ^*_{int}'s as small as .01 can inflate the variance by a factor of 1.5 over that expected without interviewer effects.[4]

---

[2] In citing previously published values of ρ^*_{int}, we note that most of these accounts provide poor descriptions of the survey design features that might affect interviewer errors—the question format and wording, interviewer training procedures, nonresponse and recording error differences across interviewers, and any other effects of the survey organization. Often only a small number of survey statistics are presented, sometimes only to illustrate the *largest* interviewer effects observed.

[3] A separate appendix describing interpenetration problems in a centralized facility is available from the authors.

Perhaps the most important weakness of the past research is the lack of attention to the stability of the ρ^*_{int} estimates. None of the studies above give guidance about whether similar effects might be expected in replications of the surveys. Given the model of interviewer effects, the precision of the ρ^*_{int} is a function of the number of interviewers used and the number of cases completed by each. Many of these studies used only a few interviewers, but the reader is not cautioned about the likelihood of different levels of interviewer effects being observed in replications of the survey.[5]

MAGNITUDES AND PRECISION OF ESTIMATED INTERVIEWER EFFECTS

In the space allowed it is impossible to describe the results of each study in detail, but Table 1 presents the range of ρ^*_{int} estimates obtained in each study, along with their mean values.[6] The average values of the ρ^*_{int}'s seem to lie below .01 over the nine surveys and to exhibit large variability over the different statistics examined. The highest average ρ^*_{int} (.0184) is found for one of the Consumer Attitude Surveys, but comparison over the other months for the same statistics shows means as low as .0057. The lowest average ρ^*_{int} (.0018) is found for a mainly factual survey on health conditions, Health in America. This survey also had the largest interviewer workload.[7]

Table 2 illustrates the importance of past studies' failure to measure the stability of ρ^*_{int} by presenting standard errors for 30 ρ^*_{int} statistics, 10 from each of three studies.[8] These standard errors are estimates of the dispersion of the distribution of ρ^*_{int}'s over replications of the surveys, with variation due to the selection of different interviewers from the population of all interviewers and from sampling variability of the respondents. This table illustrates well the range and magnitude of ρ^*_{int} values estimated throughout the nine surveys. *The standard errors for the measures of interviewer effects are often larger in size than the estimates themselves,* leading one to conclude that the values of ρ^*_{int} presented here could have been much larger *or much smaller* had different sets of interviewers been chosen to work on the surveys.

---

[4] This also implies that the inflation of variance is smaller for subclass statistics in the same survey, where *m* is smaller.

[5] Some authors base their conclusion that serious interviewer effects are present on the results of a standard *F*-test. This test is subject to several assumptions which may be violated. For example, the interviewer workloads often exhibit a great deal of variation and the variances within interviewer assignments may not be constant across interviewers.

[6] Some of the statistics are means computed on continuous variables. For discrete variables, ρ^*_{int} for the proportion choosing the modal response category is presented.

[7] This fact, together with other observations, has made us begin to consider enhancing the response error model underlying ρ^*_{int} to reflect larger interviewer variability during the initial cases completed by interviewers.

[8] In order to estimate the sampling variance of the ρ^*_{int}'s a jackknife estimator was used by dividing the sample into subsamples associated with each interviewer. Although the properties of the jackknife estimator are known in the case of a linear statistic, it has also performed well empirically for a variety of nonlinear estimators including regression and correlation coefficients (Kish and Frankel, 1972). These empirical results suggested its use in this case.

In order to relate our results to past research, these standard error estimates can be used to illustrate the largest ρ^*_{int}'s likely to be encountered with these designs. If we assume that the ρ^*_{int}'s are distributed roughly according to a *t*-distribution, we can estimate the value beyond which only 10 percent of the ρ^*_{int} values lie. These are called the "maximum" values of ρ^*_{int} in Table 2. The values of ρ^*_{int} presented in Table 1 are lower than the averages from past personal interview surveys; the *maximum values* of ρ^*_{int}'s, however, are in the general range of those estimates.

It has been assumed for some time that the measurement of interviewer effects would be enhanced by centralized telephone interviewing facilities. The ability to incorporate an interpenetrated design at a relatively small cost makes this an attractive setting for the evaluation of interviewers. However, it remains difficult to draw conclusions from the results presented here because of the imprecision of these estimates. The instability of the estimates is mainly due to the small number of interviewers employed for telephone surveys. Since the past literature in the area has generally failed to address the problem of imprecision, it is possible that many of the effects judged to be important in those studies would not be found in replications.

Remedies for Instability of Interviewer Effect Estimates

The researcher faced with unstable estimates of interviewer effects has the alternatives of ceasing their measurement, altering the survey design to improve their precision (chiefly by increasing the number of interviewers used), or increasing their precision through replication of surveys. Ceasing attempts to measure this source of survey error seems premature; the other two options deserve thorough comment.

Increasing the number of interviewers is not recommended in this case, because the hiring and training of larger numbers of interviewers may force a change in essential survey conditions that help to determine the magnitudes of interviewer effects in the survey data. If, for example, the number of interviewers were increased greatly, the quality of training or amount of supervisory attention given to each inter-

Table 2. Illustrative Values and Jackknife Estimates of Standard Errors of ρ^*_{int}

Estimated Proportions or Means	ρ^*_{int}	*S.E.*	*"Maximum"* ρ^*_{int}
HEALTH AND TELEVISION VIEWING			
Having had arthritis or rheumatism	.0227	.0132	.0400
Having had diabetes	−.0059	.0053	.0010
Not drinking beer in month	.0040	.0091	.0159
Not drinking hard liquor in month	.0152	.0117	.0305
Not visited doctor in 6 months	.0266	.0144	.0455
Reporting week of last doctor's visit	.0310	.0152	.0509
Number of ameliorative health behaviors reported	.0285	.0141	.0470
Number of ways mentioned that TV is good or bad for children	.0157	.0129	.0326
Number of "bad reactions" to medicine	.0076	.0133	.0250
Number of health symptoms reported	−.0032	.0062	.0049
HEALTH IN AMERICA			
Number of doctor visits past 2 weeks	.0092	.0076	.0192
Number of hospitalizations during the past year	.0040	.0041	.0094
Reporting excellent health status	.0085	.0077	.0186
Reporting 2–4 doctor visits past year	.0002	.0035	.0048
Mean number of chronic conditions	.0096	.0063	.0179
No bed days during the past year	−.0070	.0029	−.0032
No operations past year	−.0031	.0030	.0008
No bed days past 2 weeks	−.0057	.0025	−.0024
2 wks–6 months since doctor visit	.0003	.0042	.0058
2 wks–6 months since dental visit	.0017	.0048	.0080
SURVEY OF CONSUMER ATTITUDES, DEC '82			
Better off financially than last year	−.0232	.0171	−.0008
Saw increase in prices during the last 12 months	.0101	.0181	.0338
Percent change in prices during the last 12 months	−.0325	.0178	−.0092
Feel now a bad time to buy house	.0112	.0239	.0425
Feel now a good time to buy major household items	.0242	.0258	.0580
Expect decrease in interest rates next 12 months	.0089	.0271	.0444

Table 2. *(Continued)*

Estimated Proportions or Means	ρ^*_{int}	*S.E.*	*"Maximum"* ρ^*_{int}
Expect increase in unemployment next 12 months	−.0210	.0151	−.0012
Expect no change in their financial situation next 12 months	.0308	.0269	.0660
Expect family income to increase 15% or less next 12 months	.0546	.0385	.1050
Believe government's economic policy fair	.0508	.0361	.0981

viewer might be altered (ignoring the higher costs incurred with such a design). With departures from training guidelines interviewer variance might increase. Thus, we face the Heisenberg-like result that to measure a particular survey error well *in a single survey*, we may change its value.

Another way to obtain useful knowledge about the average levels of interviewer variability is to repeat measurement of interviewer effects over surveys. By cumulating results over surveys we hope to take advantage of the fact that averages of the same survey statistic over replications have greater stability than a single measure of it. Ideally, averages would be constructed from identical measures of the same population, using the same interviewer selection and staffing rules, and the same statistics. For example, four replications of a survey might reduce the standard errors of average ρ^*_{int}'s by half, relative to those from a single survey.[9] Rarely, however, are exactly equivalent measures obtained in several repeated studies. Lacking that, cumulation of ρ^*_{int} estimates over statistics and across surveys might be attempted within the same survey organization. This is similar to the construction of generalized sampling variance estimates in complex samples. Given the weaknesses of the other approaches, generalized interviewer effect estimates obtained by averaging over surveys are attractive.

In Figure 1 the cumulative distribution of ρ^*_{int} values is plotted for each study over all statistics for which ρ^*_{int} was calculated. It may be observed that in general the *shapes* of the distributions are related to sample size and numbers of interviewers, with the range of values of ρ^*_{int} decreasing as these increase in number. This mirrors the estimates of standard errors of ρ^*_{int}. We see that while each study yielded some

---

[9] This assumes independence in the selection of interviewers across the replications and no change in true values of interviewer effects over replications.

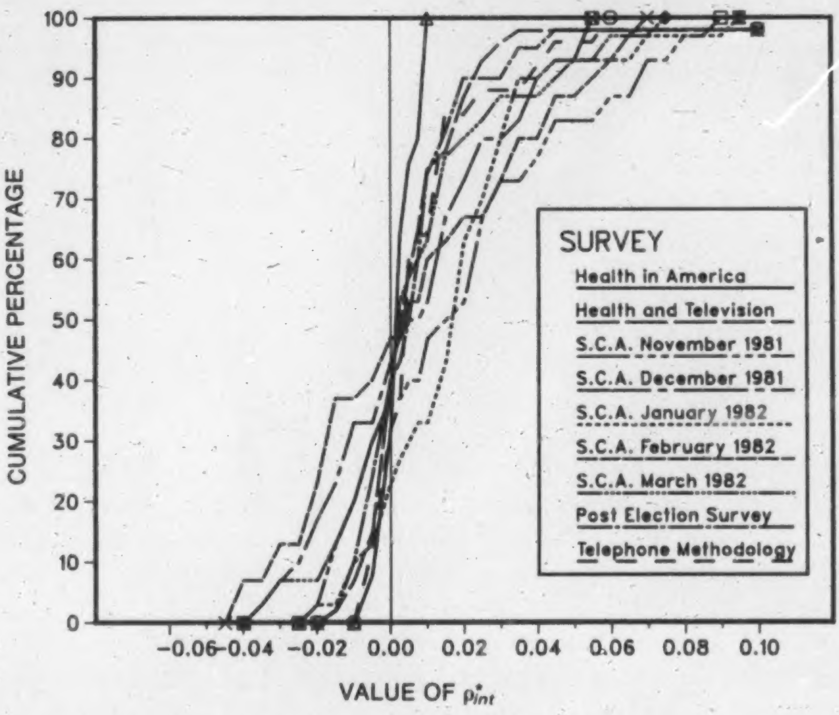

Figure 1. Cumulative Distributions of Interviewer Effect Measures (ρ^*_{int}) for Nine SRC Surveys

examples of high values of ρ^*_{int}, the majority of values cluster below 0.01. The mean values of ρ^*_{int} and the ranges are given in Table 1.

Cumulation of estimates of interviewer effects in figures such as Figure 1, together with other summary values of ρ^*_{int} over surveys and statistics, can demonstrate likely magnitudes of ρ^*_{int} to be expected, given a set of interviewer procedures. Further, as estimates of individual statistics are replicated, similar figures can be constructed for subsets of statistics (e.g., figures for questions on particular domains of attitudes; open questions versus closed questions).

The overall mean ρ^*_{int} is .009 for the nine surveys and approximately 300 statistics examined. Viewing this mean as an average of 300 independent statistics would imply a standard error about it of *approximately* .001, using the guidance of Table 2.[10] Such an average implies an increase in variance of estimates by roughly 10 percent for each 10 interviews added to the interviewers' workloads (relative to the hiring of other interviewers to complete the study). Such results might guide design decisions (following Kish, 1962) to minimize variance of statistics, given fixed resources. The optimal workload is approximately $\sqrt{Ca(1-\rho^*_{int})/c\rho^*_{int}}$, where *Ca* is the cost of hiring and training an interviewer and *c* is the cost of completing one interview. For SRC RDD surveys with 20–30-minute questionnaires, *Ca* is about $309 and *c*, about $18; thus, a ρ^*_{int} of .01 implies an optimal workload of 41 or 42 interviews.

---

[10] This is based on a weighted average sampling variance of .0004575 on individual ρ^*_{int}'s and an assumption of independence among the ρ^*_{int}'s. The lack of independence of the various ρ^*_{int}'s stems from the use of the same interviewers and respondent pairs on multiple statistics of the same survey. It is expected that the correlation among such ρ^*_{int}'s would act to increase the standard error of the mean ρ^*_{int}.

Explaining the Magnitude of Interviewer Effects

In addition to lack of concern about stability of estimates, another weakness in the interviewer effects literature is the failure to seek *explanations* for interviewer effects. The explanations for effects might lie within the nature of interviewer training guidelines, of the interviewer task (e.g., probing instructions, mode of interviewing), or of the survey instrument (e.g., question type). We were able to examine three design features—the form of the question, the type of interviewer behavior used, and the use of computer assistance in the data collection. In addition to these design features we also monitored certain aspects of interviewer behavior and examined characteristics of both the interviewers and the respondents as correlates of interviewer effects.

THE INFLUENCE OF QUESTION FORM ON INTERVIEWER EFFECTS

Previous personal interview surveys have examined question type in an effort to identify the correlates of interviewer variability. Different question types may present to the interviewer communication tasks that vary in complexity. For example, factual questions (those for which there is a knowable, verifiable answer) might be thought to be less subject to effects of different wording or inflection in their delivery than would be attitudinal questions (e.g., "How many times in the last two weeks did you visit a physician's office?" versus "How well do you like TV game shows?"). The results from past studies are mixed, some of them finding attitudinal questions subject to higher interviewer variance. Kish (1962), for example, found no important differences in interviewer effects by question type. However, several studies (Hansen et al., 1961; Fellegi, 1964; O'Muircheartaigh, 1976; Feather, 1973; Collins and Butcher, 1982) have concluded that factual items are less susceptible to interviewer effects. Larger effects have been observed for attitudinal questions, especially those with open-ended responses (O'Muircheartaigh, 1976), for emotionally charged questions (Fellegi, 1964), for difficult items such as income and occupation (Hansen et al., 1961), and for questions which lacked specification regarding the acceptable interviewing procedure (Feather, 1973).

All 297 statistics from the nine different surveys were classified as factual or attitudinal (155 factual questions and 142 attitudinal questions), and since some questions were used in several surveys, there was an opportunity for replicated comparisons. When we examine these results, there appears to be no evidence that factual items as a class are subject to any different interviewer effects than are attitudinal questions. Over all surveys, the average value of ρ^*_{int} was .0098 for factual items compared to .0085 for attitudinal items.

Another variation in question form concerns whether the respondent is supplied answer categories for his choice or is free to formulate his own answer—closed questions versus open questions. Again using data from eight of the nine different surveys, we separated open and closed questions and compared ρ^*_{int}'s. Despite suggestions that open questions might be subject to larger interviewer effects, there was no tendency for this to be the case over 63 open and 192 closed questions examined from the eight surveys. In fact, five of the eight surveys produced higher estimates of ρ^*_{int} for closed questions. Over all surveys, the average value of ρ^*_{int} for closed questions was .0082 compared to .0124 for open-ended responses.

There were, however, aspects of interviewer behavior in using open questions that did vary. For example, some open questions ask the respondent to mention several entities that belong to a certain set (e.g., "What do you think are the most important problems facing the country at this time?"); one statistic of interest is the number of different entities that are mentioned. This count is a function both of the respondent's ability to articulate problems and the interviewer's behavior in probing for more mentions. The number of responses to a question appears to be subject to greater interviewer differences than the substantive response category into which the answers are coded. For example, for six questions asked in the Surveys of Consumer Attitudes, the ρ^*_{int}'s for the percentage giving two mentions are higher each month than the ρ^*_{int} values for the modal response category of the first mention. Over all months, the average value of ρ^*_{int} for the percentage giving two mentions was .0162 compared to an average value of .0007 for the modal category of the first mention. The ρ^*_{int} for the substantive category of the second mention was .0120, close to the .0162 for the proportion with second mentions. This suggests the differential behaviors that determine whether a second mention is given also might influence substantive responses on the second mention.

EFFECTS OF INTERVIEWER BEHAVIOR GUIDELINES ON INTERVIEWER VARIABILITY

In a series of experimental surveys (of which Health in America and the study of Health and Television Viewing are two), interviewers were trained to limit their verbal interaction with the respondent to a specified set of probes or feedback, in addition to the survey questions (following Cannell et al., forthcoming). Most of these statements were entered into the questionnaire itself and were varied to fit the response behavior of the respondent. Interviewers were not permitted to engage in any other conversation with the respondent. Two groups were used: one (experimental) which limited interviewers only to feedback written in a special version of the questionnaire, and another (control) which permitted them to use at their discretion only a small number of statements in response to answers given by respondents (e.g., "I see."). Although both restricted interviewer behavior to a set of specified actions, the first technique was hypothesized to improve reporting (i.e., reduce response bias) over the second.

In both surveys the differences between the ρ^*_{int} values for the cases with specified feedback and those using feedback at the discretion of the interviewer were small relative to the standard errors of the estimates. Since one of these studies (Health in America) has unusually low ρ^*_{int}'s overall, one might speculate that both procedures reduce interviewer effects relative to typical SRC training procedures. A tendency was observed for estimates from the half-sample with specified feedback to be smaller. Combining 57 of the variables from both surveys, 58 percent of the estimates obtained from the experimental group were smaller (mean value of $-.0024$) than those calculated for the control group (mean value of .0142). Such a finding merits attempts at replication over different substantive topics.

EFFECT OF COMPUTER ASSISTED TELEPHONE INTERVIEWING (CATI)

Another experimental study used terminals to display the survey questions, route the interviewer through the questionnaire, accept, and check on the values of answers recorded by the interviewer. A CATI system thus eliminates certain types of interviewer errors. One study, Health in America, randomly assigned cases to the use of a CATI system or to the use of a paper questionnaire (Groves and Mathiowetz, 1984). Each interviewer used both systems at different times throughout the study. The same questionnaire was used on both systems. Out of the 25 variables compared, the CATI cases have lower ρ^*_{int} values on 18 of the variables, with an average value of $-.0031$ compared to an average value of .0048 for the paper questionnaire. The differences, however, were small relative to the standard errors of the ρ^*_{int}'s. Attempts to replicate these results are needed.

INTERVIEWER CHARACTERISTICS AND INTERVIEWER EFFECTS

Failure to read a question exactly as printed, inability to follow skip patterns correctly, and reading a question too fast, all are believed to contribute to errors in recorded data (Marquis and Cannell, 1969; Lansing et al., 1971). The Health in America survey monitored interviewers during the survey in order to measure these aspects of interviewer behavior. Two types of behavior, question reading and clarity and pace of question delivery, were analyzed in detail. Even with the special emphasis given to training interviewers for this study, interviewers showed significant variation in the proportion of questions read correctly for many of the variables. For 80 percent of the variables there was a significant difference between interviewers in the proportion of questions read well (i.e., correct pace, clear speech). We found, however, that these differences in interviewer delivery were not correlated with the magnitude of interviewer effects for statistics computed on these variables (e.g., the proportion choosing a particular category).

To search further for this relationship, we created scatterplots of the squared deviation of the individual interviewer's means from the overall study mean by scores from the monitoring data. We expected a positive relationship between an interviewer's squared deviations and the proportion of "incorrect" behavior. No such relationship was apparent. In both the Health in America and the Health and Television Viewing Surveys, we also examined other measures of interviewer performance such as their response rate, productivity, and supervisor evaluations. Again, we found no relationship between any one of these variables and interviewer deviations from survey averages. The so-called "better" interviewers on these characteristics did not deviate any more or any less from the overall mean than did the other interviewers.[11]

---

[11] Most telephone data yield distributions of interviewer means that contain few outliers and have relatively smooth patterns about the overall survey mean. The Health in America survey, however, had many measures where one or two interviewers were extreme outliers to the distribution. Over different statistics the identity of the outliers varied. To evaluate the impact of these extreme values we again performed the interviewer variability analysis for five statistics with high values of ρ^*_{int} (mean value of .0075), eliminating one or two outliers in each case. For each variable the new estimate of ρ^*_{int} was much smaller, with a mean value of $-.0004$ over the five variables. In other words, in that study one or two interviewers were responsible for most of the measured variability.

RESPONDENT CHARACTERISTICS AND INTERVIEWER EFFECTS

For the Health and Television Viewing Survey we examined levels of interviewer variability by sex, education, and age of respondents. The predominance of female telephone interviewers and potential differences in their interaction with female and male respondents suggested the possibility of differences in interviewer effects by sex of respondent. We found little evidence that the gender of the respondent is related to interviewer effects. Although the mean value of ρ^*_{int} over all questions was slightly higher for female respondents, further investigation by question type found mixed results. For some question types the estimates of ρ^*_{int} were larger for males; for other question types the estimates were larger for females. In all cases the differences in the estimates were small.

Although past empirical results vary (Cannell et al., 1977; Schuman and Presser, 1977), it is reasonable to argue that poorly educated respondents might be more easily affected by the behavior and status of the interviewer than more highly educated respondents. Lower education groups may seek greater help from the interviewer in answering the questions or may use the inflection of the interviewer's voice as a cue for responses to questions they find difficult. Values of ρ^*_{int}, however, do not appear to be larger for respondents with less than 12 years of education. Even when estimates of ρ^*_{int} are calculated by question type, differences by respondent education are not apparent.

Respondents were also categorized into five groups according to age. We expected larger values of ρ^*_{int} for the oldest age group. There is some evidence that telephone surveys suffer greater nonresponse among older persons (Groves and Kahn, 1979), and we sought evidence that those who do respond might be more subject to influence by the interviewer because of their own suspicion about the nature of the survey, or greater tendency to fatigue during the 30-minute interview on the telephone. Sudman and Bradburn (1973) have noted that older respondents exhibit greater response errors on questions requiring recall of factual material. In fact, the oldest age group tends to have larger ρ^*_{int} values than other age groups. Over all question types, the oldest age group had an average value of ρ^*_{int} equal to .0274 compared to an average value of .0012 for others.[12] As we noted earlier, ρ^*_{int} measures both variation in nonresponse and response errors over interviewers. These higher ρ^*_{int}'s for the elderly may relate to variation across interviewers in nonresponse rates among the elderly.

---

[12] There is also evidence that the 35–49-year-old group experiences higher ρ^*_{int}'s (mean .0099) but there is no theoretical justification for such a finding.

Summary and Speculations

The instability of estimates from any single survey is the major impediment to measuring interviewer variability. Through cumulating results over many surveys, we obtained generalized estimates of interviewer effects. The analysis showed that for SRC telephone surveys, typical ρ^*_{int} values cluster below .01, an average somewhat smaller than those reported from personal interview surveys, which average closer to .02. We argued that such average ρ^*_{int}'s are useful for guiding decisions regarding how many interviewers to use in order to minimize the variance of estimates, given fixed available resources. In addition, the cumulation of estimates over surveys within the organization could be used to provide more specific information regarding what topics or measurement types might be subject to larger or smaller ρ^*_{int}'s and thus require smaller or larger optimal workloads.

We also attempted to study correlates of interviewer effects. Our results supported others in finding that attitudinal questions are themselves not subject to greater interviewer effects than "factual" items. Open questions are not inherently subject to greater interviewer variability, but the number of distinct answers obtained to an open question appears to be sensitive to interviewer differences. These effects do not apparently influence the substance of the *first* mentioned answer, but appear to be the result of variable probing behavior, that affects both the likelihood of the respondent giving a second mention *and* the substance of that response.

Behavior may be able to be altered to reduce interviewer effects, through use of both controlled feedback and probing and of CATI systems that control the flow and questioning sequences of the interview. These are promising results that deserve further research because they offer the designer real hope of reducing this source of survey error.

In contrast, there was an unfortunate lack of success in using common indicators of performance to predict interviewer effects. Data obtained from monitoring interviewers and the traditional indicators of interviewer quality (response rate, productivity, and supervisory evaluations) seemed to be unrelated to the errors measured by between-interviewer variability. This result casts doubt on the relevance of these indicators for survey quality versus their value for measurement of cost efficiency of surveys.

Finally, we pursued a search for respondent level attributes that were related to interviewer effects. Gender and education of respondent were essentially uncorrelated to the effect, but age of the respondent was related. Older respondents were found to exhibit larger susceptibility to interviewer effects than younger respondents. Given past low response rates among the elderly in telephone surveys, this result might reflect more differential nonresponse bias across interviewers than differential response errors.

Readers of this article need to be cautious in their inference from these results. The studies examined are chiefly health and economic attitude surveys; they use interviewers hired and trained using SRC procedures; supervision, survey administration, and response rates are those achieved by SRC. Since ρ^*_{int}'s can be affected by all these attributes of survey design and practice, these results are applicable to other centralized telephone surveys to the extent their features resemble those of our nine surveys.

References

Cannell, C., L. Oksenberg, and J. Converse
1977 Experiments in Interviewing Techniques: Field Experiments in Health Reporting 1971-1977. National Center for Health Services Research.

Cannell, C.F., R.M. Groves, L.J. Magilavy, N.A. Mathiowetz, and P.V. Miller
Forthcoming "An experimental comparison of telephone and personal health surveys." Vital and Health Statistics, Series 2 Reports.

Collins, M., and B. Butcher
1982 "Interviewer and clustering effects in an attitude survey." Journal of the Market Research Society 25(1):39–58.

Feather, J.
1973 A Study of Interviewer Variance. Department of Social and Preventive Medicine, University of Saskatchewan.

Fellegi, Ivan P.
1964 "Response variance and its estimation." Journal of the American Statistical Association 59:1016–41.

Freeman J., and E.W. Butler
1976 "Some sources of interviewer variance in surveys." Public Opinion Quarterly 40:79–91.

Groves, R.M., and R.L. Kahn
1979 Surveys by Telephone. New York: Academic Press.

Groves, R.M., and N. Mathiowetz
1984 "Computer assisted telephone interviewing: effects on interviewers and respondents." Public Opinion Quarterly 48:356–69.

Hansen, M.H., W.N. Hurwitz, and M.A. Bershad
1961 "Measurement errors in censuses and surveys." Bulletin of the ISI 38(2):351–74.

Kish, L.
1962 "Studies of interviewer variance for attitudinal variables." Journal of the American Statistical Association 57:91–115.

Kish, L., and M.R. Frankel
1972 "Inferences from complex samples." Journal of the Royal Statistical Society (B) 36:1–37.

Lansing, J.B., S.B. Withey, and A.C. Wolfe
1971 Pages 219–318 in Working Papers on Survey Research in Poverty Areas. Survey Research Center, The University of Michigan.

Marquis, K.H., and C.F. Cannell
1969 A Study of Interviewer–Respondent Interaction in the Urban Employment Surveys. Survey Research Center, The University of Michigan.

O'Muircheartaigh, C.A.
1976 "Response errors in an attitudinal sample survey." Quality and Quantity 10:97–115.
O'Muircheartaigh, C.A., and A.M. Marckwardt
1980 "An assessment of the reliability of WFS data." World Fertility Survey Conference, Methodology Session No. 6.
Schuman, H., and S. Presser
1977 "Question wording as an independent variable in survey analysis." Sociological Methods and Research 6(2):151–70.
Sudman, S., and N. Bradburn
1973 "Effects of time and memory factors on response in surveys." Journal of the American Statistical Association 68:805–15.
Tucker, Clyde
1983 "Interviewer effects in telephone surveys." Public Opinion Quarterly 47:84–95.
Waksberg, J.
1978 "Sampling methods for random digit dialing." Journal of the American Statistical Association 73(361):40–46.

The Polls: Environmental Protection

JOHN M. GILLROY AND ROBERT Y. SHAPIRO

SINCE the first Earth Day and the initial peak of the ecology movement in the early 1970s (e.g., Erskine, 1972a, 1972b; Utrup, 1978) the issue of environmental protection has had its ups and downs on the political agenda and in the public's mind. But as a result of recurring problems and serious incidents—Love Canal, Times Beach, PCBs, EDB, dioxin, vinyl chloride, contaminated groundwater, air pollution, acid rain, asbestos, radon gas emissions, the disposal of nuclear/radioactive materials (and problems with nuclear power plants generally), the Superfund and other turmoil at the Environmental Protection Agency, the contamination of wild life preserves, oil spills, and especially the disaster in Bhopal—the environment has been a persistent concern, prominent in the media and magnified by worldwide environmental groups. Hazardous waste disposal and the chemical industry, in particular, with reportedly nearly 7000 accidents involving toxic chemicals in the last five years (Diamond, 1985), are widely considered as serious national problems (e.g., John, 1985). Environmental protection was a leading issue in the last New Jersey gubernatorial election, and this year's national high school debate issue concerned water quality.

Although the Reagan administration has attempted to weaken environmental controls, the American public has been leary of such governmental deregulation. Americans continue to want the government to protect them from economic and personal harm from business and industry (Shapiro and Gillroy, 1984). The evidence (some not presented below because of space limitations) shows a consistently high level of public support for environmental protection, and this is in the shadow of forced budget cuts and the possible weakening of resolve of the Environmental Protection Agency to enforce fully Congress's environmental policy. Congress has reauthorized and invigorated the Clean Water and Clean Air Acts, the Safe Drinking Water Act, the Resource Conservation and Recovery Act, the Ocean Dumping Amendments to the Marine

---

John M. Gillroy is a Congressional Fellow of the American Philosophical Association. Robert Y. Shapiro is Assistant Professor of Political Science at Columbia University and a NORC Research Associate. The authors thank Tom Smith and Sasha Heid for their comments and assistance. Support has been provided by the Columbia University Council for Research in the Social Sciences; the responsibility for analysis and interpretation is the authors'.

Public Opinion Quarterly Vol. 50:270–279 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-270/$2.50

Protection, Research and Sanctuaries Act, and most recently, the Toxic Substances Control Act (with prospects for groundwater and acid rain legislation). In contrast to the executive branch, Congress has perceived and responded to a strong public support for setting and enforcing tough pollution standards.

The survey data during the 1980s reveal substantial and in some cases clearly increasing support for environmental regulation. The year 1983 was pivotal: the Reagan administration had wanted to overhaul the major environmental laws, but it misread or failed to anticipate public opinion (Lake, 1983; Business Week/Harris Poll, 1983; Shabecoff, 1983). In response to survey questions about specific environmental problems, large segments of the public viewed nearly all of them as important *national* issues in need of significant attention, although they may rarely be perceived as *local* problems.

Support for environmental protection has been resistant thus far to energy shortages and price increases, recession, Reagan's landslides, and the high costs of environmental programs and regulation. There was little, if any, appreciable backlash in opinion. The proportion of the public wanting less regulation or less spending has remained small, with relatively little change, even as support for deregulation or "less government" rose at the the beginning of the Reagan administration (cf. Shapiro and Gillroy, 1984). The public has been reluctant to support environmentally harmful policies which might improve the economy and energy resources, with possible exceptions such as increases in strip-mining and offshore oil exploration, although citizens may waver if faced with hard choices close to home (Rosenbaum, 1985:68–70). It is evident, however, that people do not—perhaps rationally—see as mutually exclusive a healthy environment and other important needs, or they do not want to have to choose between them (Ladd, 1982; Mitchell, 1980; Keeter, 1984).

The survey data concerning the costs of environmental regulation and protection are less extensive and revealing (we cannot tell whether respondents will genuinely be willing to pay what they say, although changes in these opinions are suggestive), but this evidence is critical for learning more about the public's policy preferences when the environment's salience wanes after the latest major news event (cf. Downs, 1972). The public almost certainly knows there are significant costs—more so now than ever before—and most respondents say that they would be willing to pay some modest amount (cf. Erskine, 1972a). Of course, the problem of costs has not often been framed in surveys or elsewhere in terms of those paid directly by citizens or by polluters (subsequently paid indirectly by the former). Americans appear to think that those who pollute should pay, and prefer that the public itself contribute through higher prices rather than tax increases (Fiorina, 1973).

One dimension of environmental protection not addressed well in opinion polls is the evaluation of *risks*: the uncertainties of the scientific and technical mechanisms of both the causes of environmental problems and their solutions, and the weighing of specific prospective costs and benefits. What tradeoffs are acceptable? Which must be decided upon soon? Should the government restrict a wide range of activities until all of the long-term as well as short-term risks are assessed? Or are many of these pursuits too important for economic progress, technological advancement, or confronting future problems of scarcity or excessive government spending? How effective are certain government actions likely to be, even with the past as a guide? How much must be done and for how long? These questions may be central to the next generation of environmental issues, and they may have to be grappled with in a public dialogue.

They have evolved from the problems that first burst into the American consciousness more than 15 years ago—and remained.

The survey questions reported here were taken from the files of the Policy Change and Public Opinion Change Project at NORC (originally supported by the National Science Foundation, grant #SES-7912969 A01). These data were compiled from published and unpublished sources, and with the assistance of the Roper Center, University of Connecticut; the Louis Harris Data Center, University of North Carolina at Chapel Hill; the Inter-university Consortium for Political and Social Research, University of Michigan; and the NORC Library. Further information about these data can be obtained from the appropriate archive or survey organization. The national sample sizes are at least 1000 unless noted otherwise.

CRI — Cambridge Reports, Inc., including Cambridge Reports (1985)
HARRIS — Louis Harris and Associates
NBC — National Broadcasting Company
NYT/CBS — *The New York Times*/Columbia Broadcasting System
NORC-GSS — NORC's General Social Surveys
ROPER — Roper Organization
SRC/CPS — Survey Research Center/Center for Political Studies, University of Michigan, American National Election Studies
TRENDEX[1] — Trendex, Inc./General Electric Quarterly Index of Company Relations
YSW — Yankelovich, Skelly, and White

References

Black, Joan S.
1983 "General Electric brings the polls to life." Public Opinion, June/July, p. 27.

Business Week/Harris Poll
1983 "A call for tougher—not weaker—antipollution laws." Business Week, January 24, p. 87.

Cambridge Reports
1985 America Looks Ahead. Cambridge: Cambridge Reports, Inc.

Diamond, Stuart
1985 "U.S. toxic mishaps in chemicals put at 6,928 in 5 years." The New York Times, October 3, p. A1, D22.

Downs, Anthony
1972 "Up and down with ecology—the 'issue-attention cycle.'" The Public Interest 28:38–50.

Erskine, Hazel
1972a "The polls: pollution and its costs." Public Opinion Quarterly 36:120–35.
1972b "The polls: pollution and industry." Public Opinion Quarterly 36:263–80.

Fiorina, Morris
1973 "Public opinion and pollution: an interpretive survey." Environmental Quality Laboratory, California Institute of Technology.

John, Kenneth E.
1985 "Other voices." The Washington Post National Weekly Edition, April 15, p. 38.

---

[1] These surveys were described very briefly by Black (1983), and trend data from them were reported recently in Shapiro and Smith (1985), regrettably without further comment. Unlike most telephone surveys that are done today by the major polling organizations, which use random digit dialing (RDD) methods, the TRENDEX studies interviewed random samples taken from names listed in local telephone directories. Consequently there may be difficulties in drawing inferences from these data about mass opinion, although evaluating *trends* may pose less of a problem.

Keeter, Scott
1984 "Problematic pollution polls: validity in the measurement of public opinion on environmental issues." Political Methodology 10:267–91.

Ladd, Everett Carll
1982 "Clearing the air: public opinion and public policy on the environment." Public Opinion, February/March, pp. 16–20.

Lake, Laura M.
1983 "The environmental mandate: activists and the electorate. Political Science Quarterly 98:215–33.

Mitchell, Robert C.
1980 Public Opinion on Environmental Issues: Results of a National Public Opinion Survey. Washington, D.C.: Resources for the Future.

Rosenbaum, Walter A.
1985 Environmental Politics and Policy. Washington, D.C.: Congressional Quarterly Press.

Shabecoff, Philip
1983 "Ruckelshaus says administration misread mandate on environment." The New York Times, July 27, p. 1, 9.

Shapiro, Robert Y., and John M. Gillroy
1984 "The polls: regulation — part I," and "The polls: regulation — part II." Public Opinion Quarterly 48:531–42, 666–77.

Shapiro, Robert Y., and Tom W. Smith
1985 "The polls: Social Security." Public Opinion Quarterly 49:561–72.

Utrup, Kathryn
1978 "Environmental public opinion: trends and tradeoffs, 1969–78." Resources for the Future, Washington, D.C.

Importance of Environmental Protection

NORC–GSS/ROPER: We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount. First . . . Are we spending too much, too little, or about the right amount on . . . Improving and protecting the environment?

NORC–GSS:

	1973	*1974*	*1975*	*1976*	*1977*	*1978*	*1980*	*1982*	*1983*	*1984*	*1985*
Too little	61%	59%	53%	55%	48%	52%	48%	50%	54%	59%	56%
About right	26	27	31	31	35	33	31	32	31	33	32
Too much	7	8	10	9	11	10	15	12	8	4	08
Dk	6	7	6	5	7	5	6	6	6	4	04
										N=485	*N*=746

ROPER:

	1971	*1973*	*1974*	*1975*	*1976*	*1977*	*1978*	*1979*	*1981*	*1982*	*1983*
Too little	56%	45%	49%	48%	50%	49%	44%	45%	48%	47%	48%
About right	25	30	30	30	31	31	35	35	36	37	38
Too much	5	15	11	13	12	12	13	14	10	10	7
Dk	14	9	9	9	7	8	7	6	7	6	6

ROPER: There are also differing opinions about how far we've gone with environmental protection laws and regulations. At the present time, do you think environmental laws and regulations have gone too far, or not far enough, or have struck about the right balance?

	1973	*1974*	*1975*	*1976*	*1977*	*1979*	*1980*	*1981*	*1982*	*1983*
Too far	13%	17%	20%	15%	20%	24%	25%	21%	16%	48%
Right balance	32	44	37	35	39	36	33	38	38	30
Not far enough	34	24	31	32	27	29	33	31	37	48
Dk	21	15	12	18	14	11	10	10	9	9

CRI: In general, do you think there is too much, too little, or about the right amount of government regulation and involvement in the area of environmental protection?

	1982	*1983*	*1984*	*1985*
Too much	11%	9%	8%	10%
About right	41	34	27	28
Too little	35	44	56	54
Dk	12	14	9	8

TRENDEX: As for their effect on your way of life in the next few years, say within 10 years, how would you rate the importance of each of the following topics . . . very important, of some importance, or not particularly important?

Air pollution control measures

Important	*2/65*	*8/72*	*8/73*	*8/74*	*8/75*	*8/76*	*11/76*	*8/77*	*11/77*	*8/78*	*8/79*	*8/80*	*5/81*	*8/81*
Very	64%	80%	71%	61%	62%	61%	65%	63%	62%	65%	57%	60%	54%	62%
Some	27	15	22	27	25	27	24	27	25	26	29	30	32	27
Not	6	5	7	11	12	11	11	10	13	9	14	10	14	11
Dk/na	2	*	*	*	*	*	*	*	*	*	*	*	*	*

Water pollution control measures

Important	*2/65*	*11/65*	*8/72*	*8/73*	*8/74*	*8/75*	*8/76*	*8/77*	*8/78*	*8/79*	*8/80*	*8/81*
Very	74%	81%	90%	83%	75%	72%	70%	71%	73%	67%	72%	75%
Some	19	14	7	14	19	20	23	22	20	26	19	20
Not	4	3	2	3	6	8	7	7	6	7	9	6
Dk/na	2	2	1	*	*	1	*	*	1	1	*	*

Controlling how land is used by developers

Important	*8/74*	*8/75*	*8/76*	*8/77*	*8/78*	*8–79*	*8/80*	*8/81*
Very	75%	67%	68%	63%	65%	61%	63%	68%
Some	15	23	22	25	27	30	29	25
Not	9	9	9	11	6	8	7	6
Dk/na	1	1	1	1	1	2	1	1

* Less than 0.5%

TRENDEX: I would like to get your opinion on several areas of important *government* activities. As I read each one, please tell me if you would like the government to do more or less or do about the same as they have been on . . .

	Cleaning up air and water pollution		*Controlling hazardous waste*		*Protecting the environment*		
	8/78	*8/80*	*8/80*	*8/81*	*2/82*	*5/82*	*8/82*
More	68%	65%	84%	85%	66%	67%	68%
Same	23	26	12	12	25	24	25
Less	8	9	3	3	9	7	6
Dk/na	1	*	1	1	1	1	1

* Less than 0.5%

TRENDEX: Would you like to see *large companies* do more, do less, or do the same as they have been in connection with . . . Cleaning up air and water pollution?

	8/78	*8/79*	*8/80*	*8/81*
Do more	83%	82%	84%	84%
Do same	14	14	13	15
Do less	2	2	2	1
Dk/na	2	1	1	*

* Less than 0.5%

CRI: In recent years there's been a lot of publicity about hazardous waste disposal sites that are now causing problems. In general, do you think most of the hazardous waste disposal sites that industry and government constructed over the last 30 or 40 years are safe or unsafe?

	1983	*1984*	*1985*
Safe	11%	13%	12%
Unsafe	70	64	72
Dk	19	22	17

CRI: Let's consider for a moment the problem of hazardous wastes. Which one of the industries listed do you think is currently creating the biggest hazardous waste problem?

	1979	*1980*	*1981*	*1982*	*1983*	*1984*
Nuclear power	20	22	22	21	25	25
Automobile	8	5	7	3	1	2
Chemical	27	33	32	34	48	43
Pesticides/insectides	8	9	10	7	8	9
Other	29	25	20	21	12	15
Dk	8	6	8	13	7	7

TRENDEX: How much of a problem are hazardous wastes in your community at present—a major problem, getting to be a major problem, or not much of a problem?

	7–8/80	*8/81*
Major problem	6%	5%
Getting to be major problem	20	17
Not much of a problem	73	76
Dk/na	1	2

Environment, Energy, and the Economy

ENERGY

SRC/CPS: Present government regulations with regard to pollution and other environmental problems limit full use of some energy sources. Do you think the *government should relax environmental protection regulations* to increase the use of these energy sources, or should the *government keep environmental protection regulations unchanged* even though this may delay the production of more energy?

	1–2/80	*4/80*	*6–7/80*	*9/80*	*11–12/80*
Regulations unchanged	39%	45%	44%	45%	46%
Relax regulations	35	28	31	30	29
Relax, qualified	15	14	13	15	12
Don't know	11	13	12	10	12
	N=999	*N*=956	*N*=839	*N*=767	

ROPER: Some people say that the progress of this nation depends on an adequate supply of energy, and we have to have it even though it means taking some risks with the environment. Others say the important thing is the environment, and that it is better to risk not having enough energy than to risk spoiling our environment. Are you more on the side of adequate energy or more on the side of protecting the environment?

	9–10/73	*1/74*	*5/74*	*9–10/74*	*2/75*	*9–10/75*	*10/76*	*9/77*	*9/79*	*9–10/80*	*9/81*	*9/82*
Adequate energy	37%	47%	37%	41%	40%	40%	33%	43%	43%	45%	39%	35%
Protect environ.	37	34	42	39	35	39	44	35	38	36	40	46
Neither, no conflict	12	13	14	12	14	13	13	14	10	12	13	14
Dk	14	6	8	8	10	9	11	8	9	7	7	5

NYT/CBS: What do you think is more important—producing energy or protecting the environment?

	7/1977	7/1979
Producing energy	30%	41%
Protecting environment	51	43
No opinion	19	16

YSW: In general, do you feel that President Carter's energy program puts too much, not enough, or just the right amount of emphasis on protecting the environment?

	1977	1979[a]	1980[a]
Too much emphasis	11%	13%	12%
Right emphasis	43	31	29
Too little emphasis	36	38	45
Not sure	9	17	14

[a] Omits "protecting."

ROPER: There has been much talk about the problem of a fuel and energy shortage in this country. Here is a list of steps that have been or could be taken to conserve supplies of energy (card shown respondent).
The first is . . . Is that something you think we should do or should not do?

	Increase offshore exploration for oil reserves under the ocean				*Allow more strip mining to produce more coal supplies*			
	6/74	*6/75*	*6/76*	*6/78*	*6/74*	*6/75*	*6/76*	*6/78*
Should	68%	70%	61%	68%	46%	48%	42%	53%
Should not	18	13	18	17	37	31	34	31
Dk	14	17	22	15	17	21	23	15

	Eliminate pollution control devices on cars since they use so much more gasoline				*Relax pollution control standards so that fuels which don't meet these standards like coal, high sulphur oil, etc. can be used*			
	6/74	*6/75*	*7/76*	*6/78*	*6/74*	*6/75*	*6/76*	*6/78*
Should	45%	48%	42%	41%	39%	45%	33%	38%
Should not	46	39	44	47	50	41	52	50
Dk	9	13	13	12	11	14	15	13

ROPER: Experts say this all means we must find ways both to conserve oil and to develop new energy sources. Here are some suggestions that have been made (card shown respondent). Would you go down that list, and for each one tell me whether it is something you think should or should not be done. First, . . .

	Increase offshore exploration for oil reserves under the ocean				*Allow more strip mining to produce more coal supplies*				*Relax pollution standards so that fuels which don't meet these standards like coal, high sulphur oil, etc. can be used*			
	2/77	*3/79*	*3/80*	*3/81*	*2/77*	*3/79*	*3/80*	*3/81*	*2/77*	*3/79*	*3/80*	*3/81*
Should	79%	80%	78%	77%	57%	50%	57%	55%	35%	39%	36%	34%
Should not	10	12	14	13	26	37	33	31	51	49	53	51
Dk	11	8	7	10	17	13	11	14	13	13	10	16

CRI: Some people oppose offshore drilling for oil and gas because they fear pollution. Other people say the need for oil is so pressing that we should be willing to suffer an occasional oil spill in order to get more oil. If offshore drilling meant that a major spill would occur about once every 5 years, would you favor or oppose more offshore drilling?

	1980	1983	1985
Favor	64%	50%	49%
Oppose	22	37	37
Dk	14	14	14

NBC: Sometimes the laws designed to protect the environment make it hard to build a needed oil refinery, a new pipeline or some other energy-related facility. Should the situation arise, which do you think is more important: protecting the environment or building a needed refinery or pipeline?

	9/1979	*10/1981*[a]
Protecting environment	40%	49%
Building energy facility	47	39
Not sure	13	12

[a] The second "needed" is omitted.

ECONOMY

CRI: Which of these two statements is closer to your opinion: We must be prepared to sacrifice environmental quality for economic growth. We must sacrifice economic growth in order to preserve and protect the environment.

	1976	*1977*	*1978*	*1979*	*1981*	*1982*	*1983*	*1984*	*1985*
Sacrifice environmental quality	21%	26%	23%	32%	26%	31%	16%	27%	23%
Sacrifice economic growth	38	39	37	37	41	41	42	42	53
Dk	41	35	39	32	33	28	42	31	24

CRI: Which of these two statements is closer to your opinion? There is no relationship between economic growth and the quality of the environment—indeed, we can have more and more goods and services and also a clean world. We cannot have both economic growth and a high level of environmental quality; we must sacrifice one or the other.

	1976	*1978*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Can have growth and a clean world	50%	47%	50%	45%	49%	49%	55%	51%
Cannot have both	23	30	31	35	35	27	28	33
Dk	27	23	19	20	16	23	17	16

TRENDEX: Now let's talk about how the nation's economy might be affected by various kinds of programs. I will read you a list of programs. As I read each one, please tell me if you think it might *help* the economy, *hurt* the economy, or have little or no effect on the economy . . . An all-out attack on air and water pollution

	1–2/72	*11/75*	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*
Help	71%	58%	55%	52%	55%	53%	55%
Hurt	12	23	22	27	23	26	24
Little, no eff.	14	15	19	18	20	18	18
Dk/na	3	3	5	3	3	3	3

CRI: Would you agree or disagree: Economic policy should encourage growth in those economic areas which are less harmful to the environment and use up fewer natural resources.

	1979	*1984*
Agree	74%	81%
Disagree	10	7
Dk	16	12

ROPER: I want to ask you about some proposals that economists say would have a definite effect on controlling inflation. But while they would help in the long run, they would mean sacrifices and hardships in the next year or so. As I read each one would you tell me if you'd be willing to see it done because of its long-run benefit, or not willing to see it done because the short-term effects would be too severe? First . . . Cut back on pollution and safety regulations that raise the cost of products even if it slows progress on pollution and safety

	9–10/74	*10–11/75*	*10–11/78*	*10–11/79*
Willing	39%	39%	37%	41%
Not willing	46	45	50	46
Dk	16	16	13	13

NBC: Sometimes the laws that are designed to protect the environment cause industries to spend more money and raise their prices. If it comes to a question like that, which do you think is more important: protecting the environment, or keeping prices down?

	6/1978	*12/1978*	*10/1981*
Protect environment	60%	57%	52%
Keep prices down	31	37	37
Not sure	9	6	11

Paying Costs of Environmental Protection

TRENDEX: Would you like to see something done in your community or part of the country on the following issues: . . . If "Yes": If such action were to result in higher costs to you through taxes, utility rates or other prices, how much would you be willing to pay for . . . $5, $50, or $500?

Controlling air pollution

	1–2/66	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
Yes	81%	80%	73%	77%	80%	81%	87%	84%	72%	65%	69%	67%
No	19	20	26	21	19	19	12	15	28	34	30	32
Dk/na	*	*	*	1	1	*	1	*	1	1	1	1

	1–2/66[a]	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
$5	49%	45%	44%	43%	41%	40%	35%	36%	43%	44%	40%	41%
$50	27	28	25	30	26	30	33	34	28	29	32	33
$500	6	6	6	9	9	9	12	10	7	6	8	7
Nothing	—	8	9	5	10	9	8	9	11	11	8	8
Dk/na	18	13	15	13	14	12	12	11	11	11	12	12
(*N*)	(813)	(835)	(741)	(825)	(951)	(848)	(969)	(911)	(743)	(670)	(711)	(693)

Controlling water pollution

	1–2/66	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
Yes	90%	83%	90%	84%	89%	86%	91%	93%	91%
No	10	17	10	15	11	13	9	7	9
Dk/na	*	1	*	1	1	1	1	*	*

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
Yes	81%	84%	83%	79%	81%	77%	78%
No	18	14	16	20	19	22	21
Dk/na	1	2	1	1	*	1	1

	1–2/66[a]	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
$5	45%	48%	46%	45%	39%	45%	43%	37%	38%
$50	28	23	27	25	30	24	25	33	33
$500	7	5	5	7	8	8	11	10	11
Nothing	—	8	6	5	5	8	9	8	8
Dk/na	20	16	16	17	17	16	12	12	10
(*N*)	(924)		(935)	(849)	(939)		(949)		(973)

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
$5	43%	40%	41%	45%	42%	41%	43%
$50	27	28	29	29	33	33	31
$500	6	6	8	5	7	6	7
Nothing	8	11	9	11	8	7	6
Dk/na	16	15	13	11	10	13	13
(*N*)	(840)	(869)	(844)	(813)	(833)	(783)	(803)

Increase national parks and wilderness areas

	1–2/66	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
Yes	73%	70%	70%	69%	68%	72%	73%	78%	75%
No	26	29	30	29	30	27	26	21	23
Dk/na	1	1	*	1	2	2	1	1	1

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
Yes	68%	69%	67%	65%	68%	63%	63%
No	31	29	31	34	31	35	37
Dk/na	1	1	1	1	1	2	1

	1–2/66[a]	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
$5	45%	48%	47%	46%	43%	45%	45%	39%	39%
$50	26	21	27	23	28	22	25	31	31
$500	6	5	4	6	7	6	8	10	11
Nothing	—	9	6	5	5	9	9	9	8
Dk/na	23	16	16	20	18	19	13	12	11
(*N*	(754)		(729)	(696)	(716)	(841)	(769)	(856)	(808)

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
$5	42%	40%	41%	42%	39%	39%	41%
$50	26	30	29	32	33	33	30
$500	9	7	9	7	11	8	10
Nothing	9	8	9	9	8	8	7
Dk/na	15	15	12	9	9	12	12
(*N*)	(704)	(717)	(684)	(668)	(701)	(646)	(647)

[a] Response categories were coded differently.
* Less than 0.5%.

CRI: How much more per month would you personally be willing to pay for all the goods and services you use as a consumer if you knew that as a result of your paying higher prices business and industry would be able to operate in a way that did not harm the environment?

	1984	*1985*
None	39%	32%
$1–5	19	22
$6–10	11	14
$11–25	9	10
$26–50	4	4
Over $50	4	4
Dk	13	15

NYT/CBS: Do you agree or disagree with the following statement: Protecting the environment is so important that requirements and standards cannot be too high, and continuing environmental improvements must be made regardless of cost.

	9/1981	*9/1982 (Voters)*	*4/1983*	*1/1986*
Agree	45%	52%	58%	66%
Disagree	42	41	34	27
No opinion	13	7	8	7

CRI: Some people say that the government should give more consideration to how much new environmental regulations will cost the public. Other people say that, when it comes to protecting the environment, cost should not be a factor. Which of these views is closer to your opinion?

	1983	*1985*
Consider cost	44%	43%
Don't consider	39	43
Dk	16	13